IRE Transactions

on INFORMATION THEORY

A Journal Devoted to the Theoretical and Experimental Aspects of Information Transmission, Processing and Utilization.

Volume IT-6 **DECEMBER, 1960** **Number 5**

Published Quarterly

In This Issue

PUBLISHED BY THE

Professional Group on Information Theory

IRE Professional Group on Information Theory

The Professional Group on Information Theory is an organization, within the framework of the IRE, of members with principal professional interest in Information Theory. All members of the IRE are eligible for membership in the Group and will receive all Group publications upon payment of an annual fee of $4.00.

IRE TRANSACTIONS® on INFORMATION THEORY is published in March, June, September, and December, by the IRE for the Professional Group on Information Theory, at 1 East 79th Street, New York 21, N. Y. In addition to these regular quarterly issues, Special Issues appear from time to time. Responsibility for contents rests upon the authors and not upon the IRE, the Group, or its members. Individual copies of this issue and all available back issues, except PGIT-4, may be purchased at the following prices: IRE members (one copy) $2.25, libraries and colleges $3.25, all others $4.50.

INFORMATION THEORY

IRE Transactions on Information Theory

A Journal Devoted to the Theoretical and Experimental Aspects of Information Transmission, Processing and Utilization

Volume IT-6 | December, 1960 | Number 5

Published Quarterly

TABLE OF CONTENTS

Information Rate in a Continuous Channel for Regular-Simplex Codes*

C. A. STUTT†, MEMBER, IRE

Summary—*M* equally probable symbols may be encoded as continuous waveforms whose sample values are the coordinates of the vertices of a regular simplex in *N* dimensions. These vertices or code points, therefore, are equally spaced, and the resulting waveforms are adapted particularly to low signal-to-noise communication systems. An upper bound for the information rate in an additive Gaussian noise channel, based on the use of regular-simplex coded waveforms, has been calculated. The results indicate that for signal-to-noise energy ratios less than approximately unity and for error probabilities ranging from 10^{-2} to 10^{-8}, this rate is a sizable percentage of the ideal rate which has been derived by Shannon and Tuller.

I. Introduction

In his theory for the continuous channel, Shannon has shown that the theoretical maximum for the information rate through a channel may be approached arbitrarily closely, if the signals are encoded as noise-like waveforms of sufficient length.[1,2] Rice has developed this idea in the case of an additive Gaussian-noise channel and has made an analysis of information rates based on picking at random the coefficients in a Fourier Series representation of the signals[3] from a normal universe. His work showed that the information rate for such random signals approached the ideal or maximum rate only for extremely large numbers of symbols, in excess of 2^{40}, and the question arises whether deterministic methods of choosing signals might not be better in this respect.

A deterministic method of picking waveforms is discussed in this paper. Specifically, the sample values of M waveforms, which represent M equally probable message symbols, are taken to be the coordinates of the vertices of a regular simplex in N dimensions.[4-6] A simplex in N dimensions is bounded by $N + 1$ intersecting hyperplanes; it is the analog of a triangle, which may be put in a space of two or more dimensions, and a tetrahedron, which may be put in a space of three or more dimensions. If the edges are all of equal length, the simplex is regular. The condition that the message symbols be equally probable makes it reasonable to assign the same energy to each waveform; thus the code points [a] equally spaced with respect to each other by virtue [of] the geometry of the regular simplex, and the vect[ors] connecting them to the origin have equal lengths.

Gilbert has discussed the simplex code in a study [of] the relative efficiency of codes for signaling telepho[ne] numbers.[7] This type of code has been considered [by] Poritsky in a sphere packing problem.[8] Similar ideas [in] coding have also been employed by Kharkevitch.[9] Bas[ore] has shown that the arrangement of code points afford[ed] by the regular simplex corresponds to a minimum of [the] probability of error when the message symbols are equa[lly] probable and that the transitional probabilities decre[ase] with increasing distance between code points.[10] T[he] present discussion is concerned principally with [the] information rate for the regular-simplex code. Most [of] the mathematical details are omitted, inasmuch as th[ey] are included in a report by the author, which is genera[lly] available.[11]

II. The Vector Representations of the Waveforms

The M waveforms, which represent the message sy[m]bols, may in turn be represented by vectors in an [N-]dimensional space having projections on the coordin[ate] axes which are the sample values of the waveform[s]. If v_i denotes the ith vector representing the ith wavefo[rm,] then the M-member column matrix of signal vect[ors]

$$V \equiv \begin{bmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_M \end{bmatrix}$$

may be expressed as

$$V = XL,$$

* Received by the PGIT, March 23, 1959.

† Res. Lab., General Electric Co., Schenectady, N. Y.

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 639–645; October, 1948.

[2] C. E. Shannon, "Communication in the presence of noise," Proc. IRE, vol. 37, pp. 10–12; January, 1949.

[3] S. O. Rice, "Communication in the presence of noise—probability of error for two encoding schemes," *Bell Sys. Tech. J.*, vol. 29, pp. 60–93; January, 1950.

[4] P. H. Schoute, "Mehridimenzionale Geometre II. Teil, Die Polytope," Goschensche Verlagshandlung, Leipzig, Germany; 1905.

[5] H. S. M. Coxeter, "Regular Polytopes," Methuen & Co., Ltd., London, England; 1948.

[6] D. M. Y. Sommerville, "An Introduction to the Geometry of *N* Dimensions," Dover Publications, Inc., New York, N. Y.; 1958.

[7] E. N. Gilbert, "A comparison of signalling alphabets," *Sys. Tech. J.*, vol. 31, pp. 504–522; May, 1952.

[8] H. Poritsky, "The Distribution of Points on a Sphere Its Application to a Problem in Communication Theory," Gen Electric Res. Lab., Schenectady, N. Y., Rept. No. 57-RL-1 August, 1957.

[9] A. A. Kharkevitch, "To the theory of perfect receiver," *Ele sviaz*, vol. 10, pp. 28–34, 1956; Translation No. R-1956, U.S. D of Commerce, Office of Technical Services, Washington, D.

[10] B. L. Basore, "Regular Simplex Coding," The Dikew Corp., Scientific Rept. No. 1, Contract No. AF 19(604-40 AFCRC TN-59-165, ASTIA Doc. No. AD-213609.

[11] C. A. Stutt, "Regular Polyhedron Codes," General Ele Res. Lab., Schenectady, N. Y., Rept. No. 59-RL-2202; March, 1

here X is an M by N matrix of projections,

$$X \equiv \begin{bmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & & \\ x_{M1} & \cdots & x_{MN} \end{bmatrix}, \tag{3}$$

nd L is an N-member column matrix of unit coordinate ectors.

$$L \equiv \begin{bmatrix} 1_1 \\ \vdots \\ 1_N \end{bmatrix}. \tag{4}$$

will be convenient to normalize the v_i to unit length, e., to inscribe the regular simplex in the unit N-sphere.

Certain pertinent properties of the regular simplex are dapted to this discussion of the v_i in the following atements:

1) The sum of the M signal vectors is zero,

$$\sum_{i=1}^{M} v_i = 0; \tag{5}$$

consequently, any column sum of the projections in (3) is zero;

$$\sum_{i=1}^{M} x_{ip} = 0, \qquad p = 1, 2, \cdots, N. \tag{6}$$

2) The distance d_{ij} between the ith and jth code points is

$$d_{ij} \equiv | v_i - v_j | = \sqrt{\frac{2M}{M-1}}, \tag{7}$$

and the dot product of the two associated vectors is

$$v_i \cdot v_j = -\frac{1}{M-1}. \tag{8}$$

Because of (8) the correlation matrix, which is the matrix of dot products, is particularly simple, for all off-diagonal terms are just $-1/M - 1$:

$$VV' = \begin{bmatrix} 1 & & \left(-\frac{1}{M-1}\right) \\ & \ddots & \\ \left(-\frac{1}{M-1}\right) & & 1 \end{bmatrix} \tag{9}$$

where V' denotes the transpose of V.

3) A space of at least $M - 1$ dimensions is required for the M-Simplex. Furthermore, the spacing of code points cannot be increased by adding dimensions under the constraint of fixed length vectors, *i.e.*, fixed energy waveforms.

4) The regular simplex method of locating M points on the unit N-sphere is optimum in the sense that no other method of location can make any of the spacings of code points greater then that afforded by such equi-spacing without decreasing the spacings of other points.

An iterative procedure for constructing a simplex in N dimensions is to build onto a simplex in $N - 1$ dimensions by joining the N vertices of the $(N - 1)$-simplex to a point outside this space.[4,5] A particularly advantageous orientation for purposes of analysis of the regular simplex code is obtained if this point outside the space is taken to be the point one on the new coordinate axis. To construct a regular simplex in N dimensions 1) the $(N - 1)$-simplex must be regular; 2) it must be translated upon addition of the new point so that (6) is satisfied; and 3) it must be scaled so that it is inscribed in the unit N-sphere. To illustrate this procedure in the case $N = 3$, one starts with vectors associated with a regular 2-simplex (equilateral triangle) oriented so as to give

$$\begin{aligned} v_{12} &= -\frac{\sqrt{3}}{2} 1_1 - \frac{1}{2} 1_2, \\ v_{22} &= \frac{\sqrt{3}}{2} 1_1 - \frac{1}{2} 1_2, \\ v_{32} &= 0\, 1_1 + 1\, 1_2, \end{aligned} \tag{10}$$

where the second subscript denotes the dimensions of the space. The vectors associated with a regular 3-simplex (tetrahedron) include the unit vector in the third dimension and scaled and translated versions of the vectors from the 2-space:

$$\begin{aligned} v_{13} &= A_3 v_{12} - \frac{1}{3} 1_3 = -\frac{\sqrt{3}}{2} \frac{\sqrt{8}}{3} 1_1 - \frac{1}{2} \frac{\sqrt{8}}{3} 1_2 - \frac{1}{3} 1_3, \\ v_{23} &= A_3 v_{22} - \frac{1}{3} 1_3 = \frac{\sqrt{3}}{2} \frac{\sqrt{8}}{3} 1_1 - \frac{1}{2} \frac{\sqrt{8}}{3} 1_2 - \frac{1}{3} 1_3, \\ v_{33} &= A_3 v_{32} - \frac{1}{3} 1_3 = 0\, 1_1 + \frac{\sqrt{8}}{3} 1_2 - \frac{1}{3} 1_3, \\ v_{43} &= 0\, 1_1 + 0\, 1_2 + 1_3, \end{aligned} \tag{11}$$

where the scaling factor A_3 must be $\sqrt{8}/3$ to make the new vectors unit length.

If this procedure is continued to more dimensions, one sees that the pth column of the X-matrix of projections, $p = 1, 2, \cdots, M - 1$, is simply

$$\begin{bmatrix} x_{1p} \\ \vdots \\ x_{pp} \\ x_{p+1,p} \\ x_{p+2,p} \\ \vdots \\ x_{Mp} \end{bmatrix} = \begin{bmatrix} -\frac{1}{p} F(M, p) \\ \vdots \\ -\frac{1}{p} F(M, p) \\ F(M, p) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{12}$$

where

$$F(M, p) = \left\{ \frac{1 + \frac{1}{M - 1}}{1 + \frac{1}{p}} \right\}^{1/2} . \quad (13)$$

The vectors determined by the projections in (12) will be referred to as prototype vectors and are expressed in terms of the minimum number of dimensions. If N in (3) and (4) is greater than $M - 1$, then there are superfluous dimensions and the last $N - (M - 1)$ columns of (3) are null.

Insofar as calculations based on the central values of the correlation functions, which are given in (9), are concerned, these prototype vectors are very useful; however, the corresponding waveforms are not noise-like, and the fact that the off central values of the correlation functions are not small would produce deleterious effects in a communication system wherein timing is not known. For example, the last two vectors are

$$v_{M-1} = \frac{1}{M-1} \sqrt{M(M-2)}\, 1_{M-2} - \frac{1}{M-1} 1_{M-1},$$
$$v_M = 1_{M-1}, \quad (14)$$

where it is seen that a delay of the $(M - 1)$st waveform by one sample interval giving

v_{M-1} (shifted)

$$= \frac{1}{M-1} \sqrt{M(M-2)}\, 1_{M-1} - \frac{1}{M-1} 1_M \quad (15)$$

changes the cross correlation to $[1/(M-1)] \sqrt{M(M-2)}$, which for large M is only slightly less than the central value of unity for the auto-correlation functions. Thus, in the presence of noise, such cross-correlation peaks could not be distinguishable from auto-correlation peaks, and false or ambiguous detections would result.

In order to improve the waveforms in this respect, one may generate a new set of vectors by rotating the prototype vectors in the space, or, equivalently, project them onto a rotated set of coordinates. Thus, waveform design may be regarded geometrically as a problem of rotation. The number of used dimensions may also be increased any desired amount in the process, which would be advantageous if the noise is power limited rather than power-density limited. This rotation implies finding an orthogonal matrix A such that

$$L = AK \quad (16)$$

where K is a column matrix of k_p, $p = 1, 2, \cdots, N$, the unit coordinate vectors in the new coordinate system. Thus,

$$V = XL = XAK \equiv YK \quad (17)$$

where $Y \equiv XA$ may be identified as an M by N matrix of projections in the new coordinate system. Actually, since the last $N - (M - 1)$ columns of X are null, only $M - 1$ rows are required in the A matrix.

It is beyond the scope of this paper to discuss picking of an A matrix which will rotate the coordin system so that more suitable correlation functions obtained. However, if N is large, a random rotation m be satisfactory for some purposes; in which case an matrix which is orthogonal only in a probabilistic se may be formed by picking its members from a norm universe.

III. Information Rate

A communication system which might employ regul simplex coded waveforms is shown in Fig. 1. The wa form generator may be thought of as a tapped de line filter, for purposes of illustration, even though su a generation technique may not be practical when is very large. A particular waveform is selected for ea transmission and the received waveform is this wavefo perturbed by the additive channel noise. The recei includes a set of M filters which are individually match to the M possible transmitted waveforms. These match filters perform the operations of cross correlation wh are ideal according to the theory of Woodward a Davies.[12] The outputs of these filters are fed to a decisi device which selects that output which is both the larg of all outputs and exceeds some threshold level.

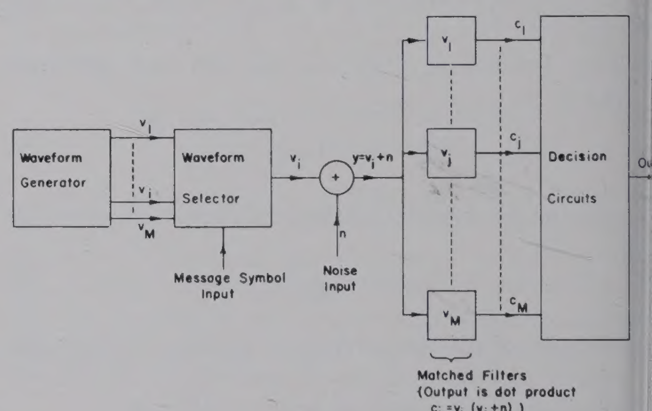


Fig. 1—Matched filter communication system employing regu simplex coded waveforms. The v_1 - - - v_M are the vector re sentations of the waveforms.

In order to determine the maximum information r for this code, attention will be focused on the cen values of the correlation functions. Thus, if $y = v_i$ + is the received vector when v_i is transmitted and n perturbing noise vector, the output of the jth filte the quantity C_j which is

$$C_j \equiv v_j \cdot (v_i + n).$$

For a correct detection when v_i is sent, then C_i m satisfy

$$C_i > C_j \qquad j = 1, 2, \cdots, i-1, i+1, \cdots, M.$$

[12] P. M. Woodward and I. L. Davies, "Information theory inverse probability in telecommunications," *Proc. IEE*, vol pt. 3, pp. 37–43; March, 1952.

d

$C_i \geq C_T$

ere C_T is the threshold level which is set to reduce the mber of false indications when no message is sent when the noncentral values of the correlation functions being computed when a message has been sent. It assumed that these noncentral values have been suitably ntrolled in the waveform design so that the number false indications arising from them is negligible.

The message symbols being equally probable, the *priori* probability for the occurrence of a particular essage vector v_i is $1/M$, so that the rate information fed into the channel is

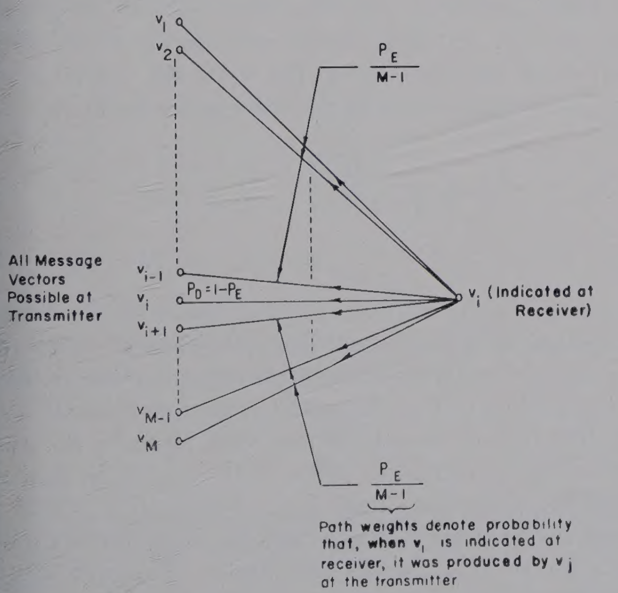
$$H(v) = \log_2 M \quad \text{bits/symbol.} \tag{19}$$

ter a decision in favor of v_i is made at the receiver, situation is described by the probability diagram of g. 2. Here the probability that v_i was sent is

$$P_E = 1 - P_E, \tag{20}$$

ere P_D is the probability of detection and P_E is the obability of error. By virtue of the symmetry of the gular simplex, the probability that v_j, $j \neq i$, was trans- tted when v_i is received is $P_E/(M - 1)$. The equivoca- n as computed from Fig. 2 is, therefore

$$(v) = -(1 - P_E) \log_2 (1 - P_E) - P_D \log_2 \frac{P_E}{M - 1} \quad \text{bits/symbol.} \tag{21}$$



. 2—Probability diagram for regular simplex code when v_i is indicated at receiver.

t is assumed that the waveforms are sent sequentially, l that no "dead" time exists between waveforms. Thus number of waveforms per unit time is just $1/T =$ $/N$, where T is the duration of the waveforms and W is their bandwidth. The actual information rate is now

$$R_a = \frac{1}{T}[H(v) - H_y(v)] \quad \text{bits/second,} \tag{22}$$

and after some mathematical manipulation,[13] a first order approximation for the case of large N is

$$R_a \approx \underbrace{W \log_2 \left\{1 + \frac{E_s}{E_n} \frac{2 \ln M}{\rho^2}\right\}}_{\text{Channel Input Terms}} - \underbrace{W \log_2 \left\{1 + \frac{E_s}{E_n} \frac{2P_E}{\rho^2} \ln \frac{e(M-1)}{P_E}\right\}}_{\text{Equivocation Terms}} \quad \text{bits/second,} \tag{23}$$

where E_s is the signal energy, E_n the noise energy in a period T and bandwidth W, and $\rho^2 \equiv N(E_s/E_n)$[14]

In general P_E will be small so that the equivocation term in (23) may frequently be neglected without serious error, so that

$$R_a \approx W \log_2 \left\{1 + \frac{E_s}{E_n} \frac{2 \ln M}{\rho^2}\right\} \quad \text{bits/second.} \tag{24}$$

If this expression is compared to the ideal information rate for a channel with additive white Gaussian noise as given by Shannon and Tuller,[1,2,15]

$$R_i = W \log_2 \left\{1 + \frac{E_s}{E_n}\right\} \quad \text{bits/second,} \tag{25}$$

it is seen that the effective signal-to-noise ratio is reduced from its actual value by the factor $(2 \ln M)/\rho^2$. It remains to determine the relationship between M and ρ^2 for prescribed P_E so that the relative values of R_a and R_i may be evaluated.

As was stated above, only the central values of the correlation functions will be considered; and it is, therefore, permissible to use the prototype vectors in the calculations, even though rotated vectors may have been used in the actual transmission. One might imagine that these transmitted vectors are given a rotation at the receiver to return them to the orientation of the prototype set. The statistics of the noise are not changed by such a rotation, and the noise vector projections (or sample values of the noise) are Gaussian before and after this transformation. It will be a further convenience to assume that the actual transmitted vector is rotated into the position of v_M [see (14)] so that it is just the unit vector 1_{M-1}. Alternatively, one might evaluate P_D and P_E for that fraction $1/M$ of the times in which v_M is transmitted; then by symmetry, the same values must hold for any other member of the code.

[13] Stutt, *op. cit.*, section V.

[14] The quantity ρ^2 is identical with the fundamental quantity R introduced by P. M. Woodward, "Probability and Information Theory," McGraw Hill Book Co., Inc. New York, N. Y., p. 87; 1953.

[15] W. G. Tuller, "Theoretical limitations on rate of information transmission," Proc. IRE, vol. 37, p. 468; May, 1949.

If $p(C_M, C_{M-1}, \cdots, C_1)$ denotes the joint density function for the matched-filter outputs, then

$$P_D = \int_{C_T}^{\infty} dC_M \int_{-\infty}^{C_M} dC_{M-1} \cdots \int_{-\infty}^{C_M} dC_1 p(C_M, C_{M-1}, \cdots, C_1). \tag{26}$$

The joint density function for the C_{M-i} may be written in terms of the conditional density functions, *i.e.*,

$$p(C_M, C_{M-1}, \cdots C_1) = p(C_M)p(C_{M-1} \mid C_M) \cdots p(C_1 \mid C_M, C_{M-1}, \cdots, C_2), \tag{27}$$

which can be written explicitly by virtue of the simple expressions for the projections, (12). Because of (5), it follows that

$$\sum_{i=0}^{M-1} C_{M-i} = 0, \tag{28}$$

so that the density function for C_1, all other C_{M-i} known, is a delta function. The remaining $M - 1$ conditional density functions are Gaussian in every case,[16] and (26) may be rewritten as

$$\begin{aligned}
P_D = {} & \frac{\rho}{\sqrt{2\pi}} \int_{C_T}^{\infty} \exp\left\{-\frac{\rho^2}{2}(C_{M-1})^2\right\} dC_M \\
& \cdot \frac{\rho}{\sqrt{2\pi}} \frac{M-1}{\sqrt{M(M-2)}} \\
& \cdot \int_{-\infty}^{C_M} \exp\left\{-\frac{\rho^2}{2}\frac{(M-1)^2}{M(M-2)}\left(C_{M-1} + \frac{C_M}{M-1}\right)^2\right\} dC_{M-1} \\
& \cdots \frac{\rho}{\sqrt{2\pi}} \sqrt{\frac{(M-1)(M-k)}{M[M-(k+1)]}} \\
& \cdot \int_{-\infty}^{C_M} \exp\left\{-\frac{\rho^2}{2}\frac{(M-1)(M-k)}{M[M-(k+1)]}\right. \\
& \cdot \left.\left(C_{M-k} + \frac{1}{M-k}\sum_{p=0}^{k-1} C_{M-p}\right)^2\right\} dC_{M-k} \\
& \cdots \frac{\rho}{\sqrt{2\pi}} \sqrt{\frac{2(M-1)}{M}} \int_{-\infty}^{C_M} \exp\left\{-\frac{\rho^2}{2}\frac{2(M-1)}{M}\right. \\
& \cdot \left.\left(C_2 + \frac{1}{2}\sum_{p=0}^{M-3} C_{M-p}\right)^2\right\} dC_2 \\
& \cdot \int_{-\infty}^{C_M} \delta\left(C_1 + \sum_{p=0}^{M-2} C_{M-p}\right) dC_1.
\end{aligned} \tag{29}$$

The integral involving C_1, denoted I_1, may be either 1 or 0, but has the expected value

$$\langle I_1 \rangle = 1 - \frac{P_E}{M-1}. \tag{30}$$

Negligible error will result for the values of P_E and M of interest, if the value of I_1 is taken as 1.

[16] Stutt, *op. cit.*, Appendix A.

If in the general integral I_k involving C_{M-k}, a chan of variable is made,

$$Z_{M-k} = \frac{\rho}{\sqrt{2}} \sqrt{\frac{(M-1)(M-k)}{M[M-(k+1)]}} \cdot \left(C_{M-k} + \frac{1}{M-k}\sum_{p=0}^{k-1} C_{M-p}\right), \qquad k = 0, 1, \cdots, M-2 \quad (3$$

then (29) may be rewritten as

$$\begin{aligned}
P_D \approx {} & \frac{1}{\sqrt{\pi}} \int_{-\rho/2(1-C_T)}^{\infty} e^{-Z^2{}_M}\, dZ_M \\
& \cdot \frac{1}{\sqrt{\pi}} \int_{-\infty}^{U_{M-1}} e^{-Z^2{}_{M-1}}\, dZ_{M-1} \\
& \cdots \frac{1}{\sqrt{\pi}} \int_{-\infty}^{U_{M-k}} e^{-Z^2{}_{M-k}}\, dZ_{M-k} \\
& \cdots \frac{1}{\sqrt{\pi}} \int_{-\infty}^{Z_2} e^{-Z^2{}_2}\, dZ_2, \qquad (3
\end{aligned}$$

where

$$U_{M-k} = \frac{\rho}{\sqrt{2}} \sqrt{\frac{(M-1)(M-k)}{M[M-(k+1)]}} \cdot \left(C_M + \frac{1}{M-k}\sum_{p=0}^{k-1} C_{M-p}\right) \qquad k = 1, \cdots, M-2. \quad (3$$

Because the U_{M-k} are random upper limits and functio of all the variables to the left of the integral I_k, it h not been possible to carry out the multiple integratio (32), exactly. An approximate solution for P_D is offer here which involves using the expected value of t U_{M-k} as an upper limit in the integral I_k, which is

$$\langle U_{M-k} \rangle = \frac{\rho}{\sqrt{2}} \sqrt{\frac{M(M-k)}{(M-1)[M-(k+1)]}} \qquad k = 1, \cdots, M-2. \quad (3$$

By using $\langle U_{M-k} \rangle$ instead of U_{M-k}, the multiple integr is replaced by a product of integrals, each of which m be evaluated in terms of the error function. The resulti approximation to P_D is proposed as an upper bound of P and this bound should become very close to the exa solution as ρ increases: Note that $\langle U_{M-k} \rangle$ is positi therefore, a fluctuation of U_{M-k} below $\langle U_{M-k} \rangle$ will c crease the value of I_k more than an equal fluctuation U_{M-k} above $\langle U_{M-k} \rangle$ will increase it. Furthermore, t variance of U_{M-k} is

$$\langle U^2_{M-k} \rangle - \langle U_{M-k} \rangle^2 = \frac{1}{2}\frac{M-(k-1)}{M-(k+1)}, \quad (3$$

which is independent of ρ; consequently, the larger t value of ρ, the further out on the tail of $e^{-Z^2{}_{M-k}}$ t integral is taken, and the effect of fluctuations of U_M on the value of I_k becomes less and less.

The value of I_k with $\langle U_{M-k} \rangle$ used as the upper limit denoted $\bar{I}_k$, and has the value

$$\frac{1}{\sqrt{\pi}} \equiv \int_{-\infty}^{\langle U_{M-k} \rangle} e^{-Z^2_{M-k}}\, dZ_{M-k}$$

$$= \frac{1}{2}\left[1 + E\left(\frac{\rho}{\sqrt{2}} \sqrt{\frac{M(M-k)}{(M-1)[M-(k+1)]}}\right)\right] \qquad (36)$$

here $E(\ )$ is the error function. The approximate expres-on for P_D may now be written as

$$_D \approx (\tfrac{1}{2})^{M-1}\left\{1 + E\left[\frac{\rho}{\sqrt{2}}(1 - C_T)\right]\right\}$$

$$\cdot \prod_{k=1}^{M-1} \left\{1 + E\left(\frac{\rho}{\sqrt{2}} \sqrt{\frac{M(M-k)}{(M-1)[M-(k+1)]}}\right)\right\}. \qquad (35)$$

or a reliable detection system, ρ will be moderately rge (perhaps in the vicinity of 10 at least) so that the rguments of the error functions are large permitting e use of the asymptotic expansion for $E(\ )$. This gives

$$P(D) \approx 1 - \epsilon_T - \sum_{k=1}^{M-2} \epsilon_k \qquad (36)$$

here ϵ_T is an error due to missed detections and is given y

$$_T = \frac{1}{2\sqrt{\pi}} \frac{\exp\left\{-\frac{\rho^2}{2}(1 - C_T)\right\}}{\frac{\rho}{\sqrt{2}}(1 - C_T)} \left\{1 - \frac{1}{\rho^2(1 - C_T)}\right\}, \qquad (37)$$

nd $\sum_{k=1}^{M-2} \epsilon_k$ is an error due to erroneous detections here ϵ_k is given by

$$= \frac{1}{2\sqrt{\pi}} \frac{\exp\left\{-\frac{\rho^2}{2} \frac{M(M-k)}{(M-1)[M-(k+1)]}\right\}}{\frac{\rho}{\sqrt{2}} \sqrt{\frac{M(M-k)}{(M-1)[M-(k+1)]}}}$$

$$\cdot \left\{1 - \frac{(M-1)[M-(k+1)]}{\rho^2 M(M-k)}\right\}. \qquad (38)$$

If the threshold is set so that ϵ_T is small compared to $\sum_{k=1}^{M-2} \epsilon_k$, an approximate expression for P_E is

$$P_E \approx \sum_{k=1}^{M-2} \epsilon_k. \qquad (39)$$

fter some mathematical manipulation,[17] an approximate xpression for P_E valid for values of M in the vicinity 10 and larger is

$$_E \approx \frac{M-1}{\rho\sqrt{2\pi}} \sqrt{\frac{M-1}{M}} \left(1 - \frac{M-1}{\rho^2 M}\right)$$

$$\cdot \exp\left\{-\frac{\rho^2}{2}\left(\frac{M}{M-1}\right)^2\right\}$$

$$- \frac{1}{\rho\sqrt{2\pi}} \exp\left\{-\frac{\rho^2}{2} \frac{M}{M-1}\right\} \sqrt{\frac{M-1}{M}}$$

$$\cdot \left[\frac{M\rho^2}{2(M-1)} - \frac{3(M-1)}{2\rho^2 M}\right] \ln \frac{2(M-1)^2}{1.78\rho^2 M}; \qquad (40)$$

and when M is very large this expression simplifies to

$$P_E \approx \frac{M}{\rho\sqrt{2\pi}}\left(1 - \frac{1}{\rho^2}\right) \exp\left\{-\frac{\rho^2}{2}\right\} \qquad (M \text{ large}). \qquad (41)$$

A straightforward calculation for the cases $M = 2$ and $M = 3$,[18] yields the following expression for P_E:

$$P_E \approx \frac{1}{\rho\sqrt{2\pi}\left(1 - \frac{1}{\rho^2}\right)} \cdot \exp\left\{-\frac{\rho^2}{2}\right\} \qquad M = 2, 3. \qquad (42)$$

In arriving at (42), only approximations for the error function for large arguments were required, and the "mean upper limit" approximation used for large values of M was not used.

Eqs. (40), (41), and (42) have been used to calculate the curves of P_E as a function of ρ^2 for fixed values of M shown in Fig. 3. These same results are shown in Fig. 4 with M as a function of ρ^2 and P_E the parameter. The curves of Fig. 4 provide the relations between M and ρ^2 for fixed P_E which are needed to complete the calculation of R_a, the actual information rate in (23). Accordingly, ratios of R_a to the ideal rate R_i, (25), have been calculated for values of P_E ranging from 10^{-2} to 10^{-8} and are plotted against $\log_{10} M$ and M in Figs. 5 and 6.

It will be noted from statement 3 in Section II that the maximum value of M for a simplex code is

$$M_{\max} = N + 1; \qquad (43)$$

hence from the definition of ρ^2,

$$M_{\max} = \rho^2 \frac{E_n}{E_s} + 1. \qquad (44)$$

The values of $M_{\max}$ for the specified values of input signal-to-noise ratio E_s/E_n are indicated by the dashed curves in Figs. 4, 5, and 6. The intersection of these dashed curves with the curves of M vs ρ^2 in Fig. 4, then, are the limits imposed by the simplex geometry on the values of M which may be used with a given input signal-to-noise ratio and a given probability of error.

These limiting values of M, as shown in the rate curves of Figs. 5 and 6, illustrate the applicability of simplex coding to low signal-to-noise ratio systems. To the right of the dashed curves, more code points should be distributed over the N-sphere than is made possible by the simplex. The curves of Fig. 5 are for values of E_s/E_n less than about 0.1. Here the dependence of R_a/R_i on E_s/E_n is negligible, so that one curve suffices for each

[17] *Ibid.*, Appendix B.

[18] *Ibid.*, Appendix C.

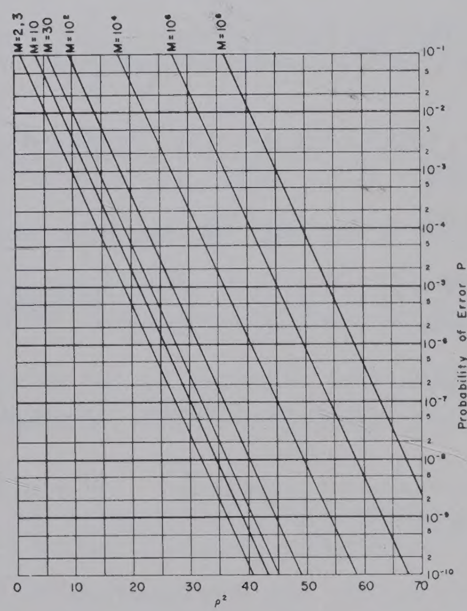


Fig. 3—Relationship between probability of error P_E and ρ^2 for fixed values of M.

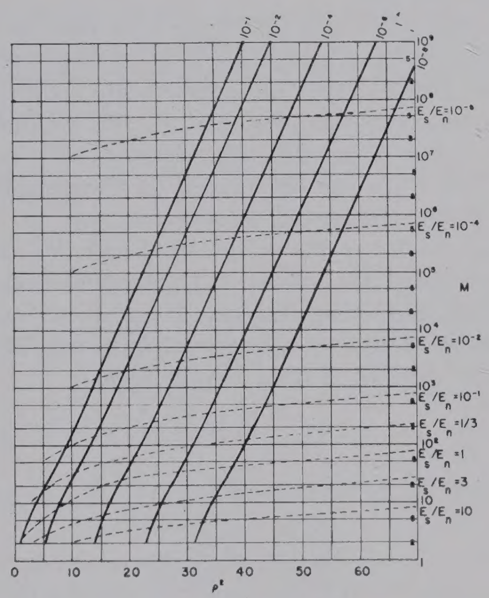


Fig. 4—Relationship between number of message symbols M and ρ^2 for fixed values of P_E. Dashed curves give maximum values of M, which are possible with regular simplex code, for the indicated input signal-to-noise ratios.

value of P_E. The curves of Fig. 6 are for values of E_s/E_n in the vicinity of 1. At these higher signal-to-noise ratios, R_a/R_i improves noticeably with E_s/E_n, hence each curve is identified by the value of both E_s/E_n and P_E. Note that except for extremely small values of P_E, the equi-spacing of code points afforded by the simplex should not be used for signal-to-noise ratios much above 10.

It would be instructive to compare the information rate curves with those obtained by Rice.[3] Such a comparison is not readily made inasmuch as Rice's curves apply for values of M much in excess of 2^{40}. However, his curve for the case of $P_E = 10^{-2}$ and $E_s/E_n = 10$ has been extrapolated to values of M under 10^3 and is

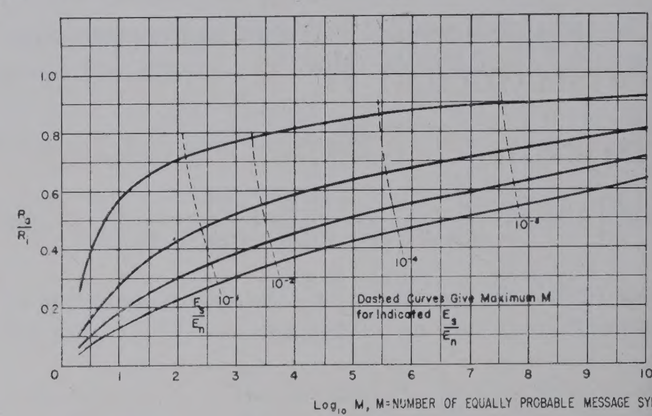


Fig. 5—Ratio of actual information rate R_A to ideal rate for addit white Gaussian noise channel R_i for regular simplex code a for low input signal-to-noise ratios (<0.1).

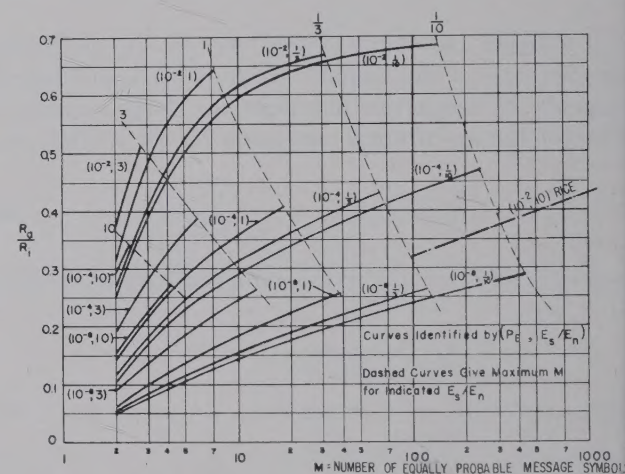


Fig. 6—Ratio of actual information rate (R_a) to ideal for addit white Gaussian noise channel (R_1) for simplex code and input signal-to-noise ratios in the vicinity of unity.

shown in Fig 6. Such a limited comparison indica that for the same conditions, the deterministic simp code does give a more rapid approach to the ideal formation rate than does a random selection.

IV. Conclusions

The regular simplex affords a convenient approach equi-spacing code points in an M-symbol continuous co Such a code is particularly applicable to low signal- noise ratio systems, and the actual rate of transmiss of information in an additive white Gaussian noise chan is a large fraction of the ideal rate given by Shann and Tuller.

Because the simplex method of coding is readily ada able to any number of symbols, from two upwards, a to an arbitrary number of dimensions ($N \geq M -$ it may be of considerable practical interest. While simplex fixes the relative position or spacing of the c points and the central values of the correlation functi the actual waveforms are arbitrary at this point. A ma problem remaining with this method of coding, then that of waveform design, and this, as has been seen a matter of suitably orienting the simplex in the N-spa

Construction of Relatively Maximal, Systematic Codes of Specified Minimum Distance From Linear Recurring Sequences of Maximal Period*

C. N. CAMPOPIANO†, MEMBER, IRE

Summary—Relative to a distance function which is both trans-tion-invariant and expressible as the sum of the distances between ordinates, an upper bound is obtained for the size of certain , d) systematic codes. This bound is closely related to a result M. Plotkin. It is shown that certain (n, d), systematic codes btainable from linear recurring sequences are of maximal size an appropriate class of systematic (n, d) codes when the distance nction is translation-invariant and the sum of the corresponding ordinate distances. The results are specialized to the Hamming stance and to the cyclic distance of C. Y. Lee. Relative to the amming distance, the results are valid for an arbitrary Galois eld $GF(q)$. For the cyclic distance, however, the results are valid nly for prime Galois fields and for $GF(4)$. Moreover, it is shown at for the latter distance, it is impossible to set up a "translation-variant, coordinate-sum" distance which is also cyclic for any nprime Galois field except $GF(4)$.

I. Introduction

IN this section we introduce basic concepts, terminology and notation, and indicate the organization of the remainder of the paper. Let $GF(q)$ be a Galois field $q = p^k$ elements where p is a prime and k is a positive teger. Let $V_n(q)$ be the set of all q^n ordered n-tuples $= (\beta_1, \cdots, \beta_n) = (\beta_i)_1^n$, with coordinates in $GF(q)$. *ote*: Unless otherwise stated, elements of $GF(q)$ will be enoted by Greek letters, while elements of $V_n(q)$ will be esignated by ordinary lower-case: $a, b, c, \cdots$. If a is $V_n(q)$ its coordinates will be represented by the cor-sponding Greek letters with appropriate subscripts; thus $= (\alpha_1, \cdots, \alpha_n) = (\alpha_i)_{i=1}^n$. Any subset of $V_n(q)$ is called *block code*,[1] or simply a *code*, of *length n*. Members of $_n(q)$ are variously called *words, vectors, points* or *sequences*. embers of a block code are called *code words, code points*, c. The *size* of a block code C, denoted by "$| C |$", is e number of code words in the code. Relative to the sual operations,

$$\alpha_1, \cdots, \alpha_n) = (\alpha\alpha_1, \cdots, \alpha\alpha_n), (\alpha_1, \alpha_2, \cdots, \alpha_n) + (\beta_1, \cdots, \beta_n) = (\alpha_1 + \beta_1, \cdots, \alpha_n + \beta_n), \quad (\alpha, \alpha_i, \beta_i, \epsilon GF(q)). \tag{1}$$

$_n(q)$ is an n-dimensional vector space over $GF(q)$. [In), "$(\alpha_1, \cdots, \alpha_n) = (\beta_1, \cdots, \beta_n)$" means "$\alpha_i = \beta_i$ r $i = 1, 2, \cdots, n$."] We write $0 = (0, 0, \cdots, 0)$ for e additive identity element of $V_n(q)$. A subset of $V_n(q)$ is called a *systematic code* if it is a linear subspace of $V_n(q)$.[2-4] A code in $V_n(q)$ is called *nontrivial* provided that for each $[i]$ between 1 and n there is some code word, say $(x_1, \cdots, x_n)$ such that $x_i \neq 0$; all other codes are called *trivial codes*. All codes discussed below are supposed to be nontrivial.

Suppose that a distance function, $d(a, b)$, is defined on all pairs a, b of $V_n(q)$. A code C in $V_n(q)$ for which $d(a, b) \geq d$ for all elements a and b of C is said to be an (n, d) code. If the size of the (n, d) code C is not exceeded by the size of any other (n, d) code, then C is said to be *maximal*. We are really interested, in the following developments, in a kind of relative maximality as follows. A systematic (n, d) code whose size is not exceeded by that of any other *systematic* (n, d) code is said to be *relatively maximal*.

In Section II, we present some general results on systematic codes some of which are valid only for a distance function of a special type (which we have called a *translation-invariant, coordinate-sum distance*). Section III summarizes relevant results from the theory of linear recurring sequences of maximal period.[5-7] From these sequences we construct, in Section III-A, codes of known minimum distance which are relatively maximal relative to a translation-invariant, coordinate-sum distance. In Section IV, these results are specialized to the Hamming distance.[2,4] For prime Galois fields, the results of Sections II, III, and III-A are applied to Lee's cyclic distance[8] in Section V. In Section VI, the results of Section V are extended to $GF(2^2)$. It is also demonstrated there that it is impossible to define a Lee cyclic distance which is a translation-invariant, coordinate-sum distance for any nonprime Galois field other than $GF(4)$. Finally,

* Received by the PGIT, August 14, 1959.
† Airborne Systems Div., Radio Corp. of America, Camden 2, J.

[1] S. P. Lloyd, "Binary block coding," *Bell Sys. Tech. J.*, vol. , pp. 517–535; March, 1957.

[2] R. W. Hamming, "Error detecting and error correcting codes," *Bell Sys. Tech. J.*, vol. 29, pp. 147–160; April, 1950.

[3] D. Slepian, "A class of binary signaling alphabets," *Bell Sys. Tech. J.*. vol. 35, pp. 203–234; January, 1956.

[4] B. M. Dwork and R. M. Heller, "Results of a geometric approach to the theory and construction of nonbinary, multiple error and failure correcting codes," 1959 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 123–129.

[5] N. Zierler, "Linear recurring sequences," *J. Ind. and Appl. Math.*, vol. 7, pp. 31–48; March, 1959.

[6] S. W. Golomb, "Sequences with Randomness Properties," The Glenn L. Martin Co., Baltimore, Md., Terminal Progress Rept.; June 14, 1955.

[7] D. A. Huffman, "The synthesis of linear sequential coding networks," *Third London Symp. on Information Theory*, London, Eng., Butterworths Scientific Publications, London, Eng., Colin Cherry, Ed., pp. 77–95; 1956.

[8] C. Y. Lee, "Some properties of nonbinary error-correcting codes," IRE TRANS. ON INFORMATION THEORY, vol. IT-4, pp. 77–82; June, 1958.

Section VII contains a single example of the preceding analyses for the special case $GF(4)$.

Before proceeding to the developments outlined above, we add a few remarks concerning the definition of the term *systematic code* given above. We could have generalized the concept of systematic code given by Hamming[2] and Slepian[3] as follows:

Definition 2: A systematic code in $V_n(q)$ is a set of elements $a = (\alpha_i)_1^n$, $b = (\beta_i)_1^n$, $\cdots$ in $V_n(q)$ determined by the following rule. For a certain fixed set of k indices, say $i_1, \cdots, i_k$, the elements $\alpha_{i_1}, \cdots, \alpha_{i_k}$, called *information digits*, are assigned arbitrary values from $GF(q)$, while each of the remaining $n - k$ elements, say $\alpha_{i_{k+1}}, \cdots, \alpha_{i_n}$, *called check digits*, is a fixed linear combination of the information digits, say

$$\alpha_m = \sum_{j=1}^{k} \alpha_{m,j}\alpha_{ij};\ m\epsilon\{i_{k+1}, \cdots, i_n\}.$$

The k-tuple $(x_{i_1}, \cdots, x_{i_k})$ assumes each of the q^k values in $V_k(q)$.

But then such a code is necessarily a linear subspace of $V_n(q)$. Moreover, one can readily modify Slepian's proof[3] of his Theorem 3 to show that every linear subspace of $V_n(q)$ is a systematic code in $V_n(q)$ according to definition 2. It is an interesting fact that *every* additive subgroup of $V_n(q)$ is a linear subspace of $V_n(q)$ if and only if q is a prime. This explains the difference between Slepian's terminology in his Theorem 3 and our terminology above which stresses linear subspaces. On the other hand, generalization of Slepian's decoding scheme[3] to certain symmetric, nonbinary, discrete channels without memory utilizes the group concept alone.

II. General Results for Systematic Codes

In the present section, we generalize a result of Slepian's.[9] This result is applied to the case in which a special kind of distance function is defined on $V_n(q)$ to obtain a partial generalization of a result of Plotkin[10] on the maximal size of systematic (n, d) codes. For nontrivial codes we have the following result.

Theorem 1: Let C be a nontrivial code in $V_n(q)$ of size $|C| = q^m$ $(1 \leq m \leq n)$. Let the code words of C be $(c_{i1}, \cdots, c_{in})$, $i = 1, 2, \cdots, q^m$. Then in the set $\{c_{1j}, c_{2j}, c_{3j}, \cdots, c_{q^m j}\}$ $(j = 1, 2, \cdots, n)$, each of the q elements of $GF(q)$ occurs precisely $q^{m-1} = (1/q)\,|C|$ times. Hence, in the code C each of the q elements of $GF(q)$ occurs precisely $(n/q)\cdot|C|$ times.

Proof: For each ξ in $GF(q)$, let N_ξ be the number of code words in C whose first coordinate is ξ. In particular, N_0 is the number of code words in C whose first coordinate is 0, where 0 is the additive identity of $GF(q)$. We shall show that N_ξ is the same for all ξ, hence has the value q^{m-1}. Let α be an arbitrary nonzero element of GF There exists some code word, say $c = (\gamma_1, \gamma_2, \cdots,$ such that $\gamma_1 \neq 0$. Hence there exists at least one co word with α as its first coordinate, namely $\alpha\gamma_1^{-1}c$. That $N_\alpha > 0$. Let a_i $(i = 1, 2, \cdots, N_\alpha)$ be the N_α disti code words with α as first coordinate. Then $a_1 -$ $a_2 - a_1, a_3 - a_1, \cdots, a_N - a_1$ are N_α *distinct* code wo with 0 as first coordinate. Hence $N_0 \geqq N_\alpha$. Let $Z_2, \cdots, Z_{N_0}$ be the distinct code words with 0 as fi coordinate. Then, since a_1 has α as first coordina $Z_i + a_1$, $i = 1, 2, \cdots, N_0$ are *distinct* code words w α as first coordinate. Hence $N_\alpha \geq N_0$ and $N_0 =$ But α was an *arbitrary* nonzero element of $GF(q)$. He for all $\xi \neq 0$, $N_\xi = N_0 = q^{m-1}$. The remainder of proof is clear.

Corollary 1: Let $F(a)$ be a real-valued function defi on $V_n(q)$ by the relation

$$F(a) = \sum_{i=1}^{n} f(\alpha_i) \qquad a = (\alpha_1, \cdots, \alpha_n),$$

where $f(\alpha)$ is a real-valued function defined on $GF($ Let C be a nontrivial code in $V_n(q)$. Then

$$\sum_{a\epsilon C} F(a) = \frac{n}{q}\,|C|\cdot \sum_{\alpha\epsilon GF(q)} f(\alpha).$$

The above results are particularly significant in situ tions where we have a distance function of a spec type which we have called a *translation-invariant*, *ordinate-sum* distance. (We are not pleased with t terminology, but some simple labels seemed appropri in order to underline the specialized nature of the f lowing results.) We introduce a translation-invaria coordinate-sum distance on $V_n(q)$ as follows: Let th be defined for each α of $GF(q)$ a *norm* or *weight* $|\alpha|$ su that $|\alpha|$ is real-valued and

$$(n_1^0)\ |-\alpha| = |\alpha|;\ |0| = 0,$$

$$(n_2)\alpha \neq 0 \quad \text{implies} \quad |\alpha| \neq 0,$$

$$(n_3)\ |\alpha + \beta| \leq |\alpha| + |\beta|.$$

Then

$$(n_4)\ |\alpha| \geq 0,$$

and the function $\rho(\alpha, \beta) = |\alpha - \beta|$ is a distance funct on $GF(q)$. Perhaps we should call $|\alpha|$ a "weak" no since we do not demand that $|\alpha\beta| = |\alpha|\,|\beta|$.[11] Ho ever, we shall simply say "norm." We define the *no* or *weight* of any element, $b = (\beta_1, \cdots, \beta_n)$ of $V_n(q)$

$$||b|| = \sum_{i=1}^{n} |\beta_i|.$$

Then $||b||$ is real-valued and satisfies $(n_1^0), \cdots,$ (when "| |" is replaced by "|| ||". We define the dista $d(a, b)$, between two elements $a = (\alpha_1, \cdots, \alpha_n)$ b

[9] D. Slepian, *op. cit.*, p. 228.

[10] M. Plotkin, "Binary Codes with Specified Minimum Distance," M. S. thesis, Moore School of Elec. Engrg., University of Pennsylvania, Philadelphia, Pa.; June, 1952.

[11] M. H. A. Newman, "Elements of the Topology of Plain of Points," Cambridge University Press, Cambridge, Eng.; 1

$\beta_1, \cdots, \beta_n)$ of $V_n(q)$ as $d(a, b) = || a - b ||$. Clearly,

$$d(a, b) = \sum_{i=1}^{n} \rho(\alpha_i, \beta_i) = \sum_{i=1}^{n} | \alpha_i - \beta_i |. \tag{3}$$

/e call $d(a, b)$ a translation-invariant, coordinate-sum istance because of (3) and because $d(a, b)$ is translation-ıvariant ($d(a + c, b + c) = d(a, b)$). Note that $\rho(\alpha, \beta)$ also translation-invariant in that

$$\rho(\alpha + \gamma, \beta + \gamma) = \rho(\alpha, \beta). \tag{4}$$

In the above, we started with $| \alpha |$, defined $\rho(\alpha, \beta) =$ $\alpha - \beta |$ and continued to define $|| b ||$, etc. Alternatively, iven a distance function, say $\rho(\alpha, \beta)$, for $GF(q)$ such hat $\rho(\alpha, \beta)$ is translation-invariant, we can define $| \alpha | =$ $(\alpha, 0)$. Then $| \alpha |$ satisfies $(n_1^0), \cdots, (n_4)$ so that it is genuine norm as defined above. The definitions for $b ||$ and $d(a, b)$ can then proceed as above. Alternatively, re could postulate a translation-invariant distance function $d(a, b)$ for $V_n(q)$ and a function $\rho(\alpha, \beta)$ such that $(a, b) = \sum_{i=1}^{n} \rho(\alpha_i, \beta_i)$. Then $\rho(\alpha, \beta)$ is a translation-ıvariant distance function for $GF(q)$, and the functions $(a, 0)$, $\rho(\alpha, 0)$ play the roles of the norms $|| a ||$, $| \alpha |$ ıtroduced above.

Let $S_q(n, d)$ represent the size of the largest nontrival, ystematic (n, d) code in $V_n(q)$. From the above definitions nd Corollary 1 we readily obtain Theorem 2.

Theorem 2: If

$$d > \frac{n}{q} \cdot \sum_{\alpha \varepsilon GF(q)} | \alpha |,$$

hen

$$S_q(n, d) \leq \frac{d}{d - \frac{n}{q} \sum_{\alpha \varepsilon GF(q)} | \alpha |}$$

vhenever the distance function, $d(a, b)$, is a translation-ıvariant, coordinate-sum distance function defined in erms of $| \alpha |$ as in (3).

Proof: Let C be a nontrivial (n, d) code in $V_n(q)$. Then $d(a, b) \geqq d$ for all a, b in C. In particular, $d \leq$ $(a, 0) \equiv || a ||$ for all nonzero a in C. Hence, since there re $(| C | - 1)$ nonzero terms in C, we have

$$\sum_{a \varepsilon C} || a || \geqq (| C | - 1)\, d.$$

Therefore, from Corollary 1, and (2),

$$\frac{n}{q} \cdot | C | \sum_{\alpha \varepsilon GF(q)} | \alpha | \geqq (| C | - 1)\, d.$$

Then the theorem follows directly.

III. Linear Recurring Sequences of Maximal Period

In this section we utilize the notation and terminolgy of Zierler[5] in summarizing certain results from the heory of linear recurring sequences which are required the following sections. We consider linear recurring sequences over the Galois field $GF(q)$, $q = p^k$. Recall that a linear recurring sequence of mth order is a sequence $\{\alpha_i\}_1^\infty$ which satisfies a linear recurrence relation, *i.e.*, a relation of the form

$$\gamma_0\alpha_k + \gamma_1\alpha_{k-1} + \cdots + \gamma_m\alpha_{k-m} = 0, \qquad k = m + 1, m + 2, \cdots \tag{5}$$

with $\alpha_i, \gamma_i, \varepsilon\, GF(q)$ and $\gamma_0\gamma_m \neq 0$. The mth order recurring sequence is uniquely determined by the relation (5) and by the first m terms $\alpha_1, \alpha_2, \cdots, \alpha_m$. It is known that these sequences are periodic with periods not exceeding $q^m - 1$; that is, there exists a $T \leq q^m - 1$ such that $\alpha_{i+T} = \alpha_i$, $i = 1, 2, \cdots$. There are exactly $\phi(q^m - 1)/m$ translation-distinct linear recurring sequences of order m with period $q^m - 1$ [$\phi(s)$ is the number of integers in the set $1, 2, \cdots, s$ which are relatively prime to s]. Following Zierler we call such a sequence with the maximal period an M *sequence*. We shall use M sequences to construct systematic, relatively maximal (n, d) codes. First of all, however, we recall a few facts about M sequences and their recurrence relations.

The *characteristic polynomial* $f(x)$ of the relation (1) is defined as

$$f(x) = \gamma_m x^m + \gamma_{m-1}x^{m-1} + \cdots + \gamma_1 x + \gamma_0. \tag{6}$$

It is known that a necessary (though not sufficient) condition that the nonzero sequences determined by (5) be M sequences is that the characteristic polynomial (6) be irreducible in the ring, $GF(q)[x]$, of polynomials with coefficients in $GF(q)$. Some irreducible polynomials are given for $GF(p)$, $p = 2, 3, 5$ and 7 by Church.[12] A few polynomials which are irreducible over $GF(2^2)$ are displayed in Section VII of this paper.

Let $f(x)$ be an irreducible polynomial which is the characteristic polynomial of an mth order linear recurrence (5) which yields M sequences. Let $G(f)$ denote the set of all sequences generated by (5). There are q^m elements in $G(f)$ including the *zero* sequence, $\phi = (0, 0, \cdots)$. All $q^m - 1$ nonzero sequences of $G(f)$ are M sequences of period $q^m - 1$. The M sequences of $G(f)$ differ from each other merely by a translation; *i.e.*, if $\{\alpha_i\}_1^\infty$, $\{\beta_i\}_1^\infty \varepsilon\, G(f)$ and are nonzero, then there exists an integer s such that $\beta_{i+s} \equiv \alpha_i$, $i = 1, 2, 3, \cdots$. If $\{\alpha_i\}_1^\infty$ is an M sequence of $G(f)$, then every subsequence $\{\alpha_i\}_s^{s+q^m+m-2}$ has the property that each of the $(q^m - 1)$ ordered in m tuples $(\beta_1, \cdots, \beta_m)$ with elements in $GF(q)$ (exclusive of the all-zero m-tuple) occurs in the subsequence precisely once. It is known that when $G(f)$ contains M sequences then $G(f)$ is an m-dimensional vector space relative to the operations

$$\alpha\{\alpha_i\}_1^\infty = \{\alpha\alpha_i\}_1^\infty$$
$$\{\alpha_i\}_1^\infty + \{\beta_i\}_1 = \{\alpha_i + \beta_i\}_1^\infty \tag{7}$$

[12] R. Church, "Tables of irreducible polynomials for the first four prime moduli," *Annals of Math.*, vol. 36, pp. 198–209; January, 1935.

for $\{\alpha_i\}_1^\infty$, $\{\beta_i\}_1^\infty \varepsilon G(f)$, $\alpha \varepsilon GF(q)$. The preceding well-known results yield directly the following.

Theorem 3: Let $\{\alpha_i\}_1^\infty$ be an mth order M-sequence (with period $q^m - 1$) over the field $GF(q)$. Define the set of $(q^m - 1)$-tuples $a_0, a_1, \cdots, a_{q^m-1}$ as

$$a_0 = (0, 0, \cdots, 0)$$
$$a_i = (\alpha_i, \alpha_{i+1}, \cdots, \alpha_{i+q^m-2}), i = 1, 2, \cdots, q^m - 1. \tag{8}$$

Then for $i \circ j \neq 0$, a_i and a_j are cyclic permutations of each other. If a_i is in the set $a_1, \cdots, a_{q^m-1}$, then 0 occurs exactly $q^{m-1} - 1$ times in a_i while each of the $(q - 1)$ nonzero elements of $GF(q)$ occurs precisely q^{m-1} times in a_i. The set (8) is an m-dimensional linear subspace of $V_{q^m-1}(q)$.

A. Application to the Construction of (n, d) Codes

Let a translation-invariant, coordinate-sum distance $d(a, b)$ be defined for $V_n(q)$ in terms of the norms $|\alpha|$, $||a||$ and the distance function $\rho(\alpha, \beta)$ as indicated in Section II. The results of Sections II and III then yield Theorem 4.

Theorem 4: The set (8) is a systematic code of size q^m in $V_n(q)$ $(n = q^m - 1)$. Relative to a translation-invariant, coordinate-sum distance, the distance between code words is $d = q^{m-1}N$, where

$$N = \sum_{\alpha \varepsilon GF(q)} |\alpha|.$$

The code (8) is a relatively maximal $(q^m - 1, q^{m-1}N)$ code. In fact,

$$S_q(q^m - 1, q^{m-1}N) = q^m.$$

Proof: For $a_i \neq a_j$ in (8), $d(a_i, a_j) = ||a_i - a_j|| = ||a_1|| = q^{m-1}N = d$. Moreover,

$$d - \frac{n}{q} N = \frac{1}{q} N > 0.$$

Hence, from Theorem 2

$$S_q(n, d) \leq \frac{q^{m-1}N}{\frac{1}{q} N} = q^m.$$

But the size of code (8) is q^m. Hence $S_q(n, d) = q^m$ and (8) is indeed relatively maximal.

IV. Application to the Hamming Distance

We utilize the definition of the (generalized) Hamming distance as given in Dwork and Heller,[4] but we introduce it in terms of the concepts used in Section II. Thus, we define the *Hamming norm* of elements, α, of $GF(q)$ as $|\alpha| = 1$ or 0 according as α is nonzero or zero. The *Hamming norm* of any $b = (\beta_1, \cdots, \beta_n)$ in $V_n(q)$, namely $||b||$, is defined as in Section II, *i.e.*, $||b|| = \sum_{i=1}^n |\beta_i|$. Finally the Hamming distance, $d(a, b)$, between any elements a, b in $V_n(q)$ is $d(a, b) = ||a - b||$. Then, clearl

$$\sum_{\alpha \varepsilon GF(q)} |\alpha| = (q - 1).$$

Hence the following results are immediately evident.

Theorem 5: Relative to the Hamming distance a the Hamming norm, the following results are valid:

1) The sum of the norms of all the code words in systematic, nontrivial code C in $V_n(q)$ is $\sum_{a \varepsilon C} ||a||$ $n/q \cdot |C| \cdot (q - 1)$.

2) If $d > (n/q)(q - 1)$, then

$$S_q(n, d) \leq \frac{d}{d - \frac{n}{q}(q - 1)}.$$

Theorem 6: Relative to the Hamming norm a metric, the set (8) is a systematic code of size q^m a length $q^m - 1$ such that the distance between code wor is $q^{m-1}(q - 1)$. This code is relatively maxima $S_q(q^m - 1, q^{m-1}(q - 1)) = q^m$.

V. Applications to the Cyclic Distance for Prime Fields of Odd Order

We utilize the cyclic distance of Lee.[8] Following Le we can order the elements of $GF(q)$ on a circle, sep rating each pair of adjacent elements by a unit a and define the distance, $\rho(\alpha, \beta)$, between pairs of fie elements α, β as the minimum number of arcs betwe α and β. Then the distance between $a = (\alpha_i)_1^n$ and b $(\beta_i)_1^n$ is $\delta(a, b) = \sum_{i=1}^n \rho(\alpha_i, \beta_i)$. In order to utilize t results of Sections II, III and III-A for this distanc it is necessary that $\rho(\alpha, \beta)$ be translation-invariant relati to the addition operation of $GF(q)$. Since the additi of $GF(q)$ is already fixed, we can make $\rho(\alpha, \beta)$ translatio invariant only by properly ordering the elements $\beta, \cdots$ on the circle. In this section we shall show ho to do this for prime Galois fields. In the following secti we shall give the appropriate ordering for $GF(2^2)$ a shall prove the impossibility of finding a satisfacto ordering for other nonprime fields.

We define a translation-invariant cyclic distance $GF(p)$ (p odd) as follows. First of all, we represent $GF($ by the "absolutely least" complete system of residu modulo p in the ring of integers:

$$GF(p) = \left\{0, \pm 1, \cdots, \pm \frac{p - 1}{2}\right\}$$

with the usual operations for addition and multiplicati We then define the *cyclic norm* of α in $GF(p)$, $|\alpha|$,

$$|\alpha| = \text{the absolute value of the real integer } \alpha.$$

The *cyclic distance*, $\rho(\alpha, \beta) = |\alpha - \beta|$ is then translati invariant. For elements of $V_n(p)$, the corresponding *cyc norm* and *cyclic distance* are $||b|| = \sum_{i=1}^n |\beta_i|$, $\delta(a, b)$ $||a - b||$. We then have

$$\sum_{\alpha \varepsilon GF(p)} |\alpha| = \tfrac{1}{4}(p^2 - 1).$$

ence, from Sections II, III, and III-A we readily obtain heorem 7.

Theorem 7: Relative to the cyclic norm and cyclic stance, the following are valid:

1) The sum of the norms of the code words in a nonivial systematic code C in $V_n(p)$ is

$$\sum_{a \in C} \| a \| = \binom{n}{p} \mid C \mid \cdot \frac{(p^2-1)}{4}.$$

2) If $d > \frac{n}{p} \cdot \frac{(p^2-1)}{4}$, then

$$S_p(n, d) \leq \frac{d}{d - \frac{n}{p} \frac{(p^2-1)}{4}}.$$

Theorem 8: Relative to the cyclic norm and distance, e set (8), with $q = p$, is a systematic code of size p^m ngth $p^m - 1$ such that the distance between code words $p^{m-1}(p^2-1)/4$. Moreover, (8) is relatively maximal; $_p(p^m - 1, p^{m-1}(p^2-1)/4) = p^m$.

VI. Application to the Cyclic Distance For Nonprime Galois Fields

We first carry out the program of Sections IV and V r $GF(4)$; then we show that this same program can ot be applied to other nonprime Galois fields. Consider $F(2^2)$. Let its elements be 0, 1, ω, ω^2 where 0 and 1 e respectively the additive and multiplicative identity ements of $GF(2^2)$. Special values of operation for this ld are: $1 + \omega + \omega^2 = 0$, $\omega^3 = 1$, $\alpha + \alpha = 0$ for all α his last rule is a property of all $GF(2^k)$. Starting with 0, e can place the elements of $GF(2^2)$ on a circle in the ockwise direction in the order 0, ω, 1, ω^2 and proceed in the first paragraph of Section V to define $\rho(\alpha, \beta), \cdots$. stead, we proceed directly as follows. Define a norm $GF(4)$ as $|0| = 0$, $|\omega| = |\omega^2| = 1$, $|1| = 2$, and e distance $\rho(\alpha, \beta) = |\alpha - \beta|$. Clearly $\rho(\alpha, \beta)$ is transtion-invariant. Hence, defining the norm and distance r $V_n(q)$ as $\| b \| = \sum_{i=1}^{n} |\beta_i|$, $\delta(a, b) = \| a - b \|$, e obtain

$$\sum_{\alpha \in GF(4)} |\alpha| = 4,$$

d the following, by now evident, result.

Theorem 9: Relative to the cyclic norm and cyclic stance, the following are valid:

1) The sum of the norms of the code words in a nontrivial systematic code C in $V_n(4)$ is

$$\sum_{a \in C} \| a \| = \frac{n}{4} \cdot \mid C \mid \cdot 4 = n \mid C \mid.$$

2) If $d > n/4 .4 = n$, then

$$S_4(n, d) \leq \frac{d}{d-n}.$$

Theorem 10: Relative to the cyclic norm and distance, the set (8), for $q = 2^2$, is a systematic code of size 4^m and length $4^m - 1$ such that the distance between code words is 4^m. Moreover, (8) is maximal; $S_4(4^m - 1, 4^m) = 4^m$.

We now consider a specific case for which the preceding process is impossible. Consider $GF(3^2)$, and designate its elements by $\{0 = \alpha_0, \alpha_1, \alpha_{-1}, \cdots, \alpha_4, \alpha_{-4}\}$. Starting with 0, place these elements around a circle in the clockwise direction in the following order: 0, α_1, α_2, α_3, α_4, α_{-4}, α_{-3}, α_{-2}, α_{-1}, separating adjacent elements by a unit arc. Then the distance $\rho(\alpha, \beta)$ between elements α, β of $GF(9)$ is defined as the minimum number of arcs between α and β. We shall show that $\rho(\alpha, \beta)$ can not be translation-invariant; that is, it is impossible that for all γ, $\rho(\alpha + \gamma, \beta + \gamma) = \rho(\alpha, \beta)$ where $\alpha + \gamma$ is, of course, the field sum. In fact, suppose that $\rho(\alpha, \beta)$ is translation-invariant; then we can define the norm of α as $|\alpha| = \rho(\alpha, 0)$. Then $|\alpha|$ satisfies (n_1^0), (n_2), (n_3) and (n_4) of Section II, and $\rho(\alpha, \beta) = |\beta - \alpha| = |\alpha - \beta|$. Clearly, we must have α_i equal to the absolute value of the integer i. The requirement that $|\alpha_1| = |-\alpha_1| = 1$, together with the fact that the only two elements with norm 1 are α_1 and α_{-1}, implies that either $-\alpha_1 = \alpha_1$ or $-\alpha_1 = \alpha_{-1}$. But $2\alpha_1 \neq 0$. Hence $\alpha_{-1} = -\alpha_1$. Similarly, $-\alpha_i = \alpha_{-i}$ for $i = 2, 3, 4$. The distance between α_1 and α_2 is 1. Hence we must have $|\alpha_2 - \alpha_1| = 1$; hence $\alpha_2 - \alpha_1 = \pm\alpha_1$. But $\alpha_2 - \alpha_1 = -\alpha_1$ implies $\alpha_2 = 0$, so that we must have $\alpha_2 - \alpha_1 = \alpha_1$. Therefore, $\alpha_2 = 2\alpha_1$. Similarly, we can show that $\alpha_4 = 2\alpha_2$ so that $\alpha_4 = 4\alpha_1 = \alpha_1$. Then $|\alpha_4| = |\alpha_1|$; but this is obviously impossible. Therefore $\rho(\alpha, \beta)$ is not translation-invariant. The ordering of the elements of $GF(9)$ on the circle was arbitrary, since the one-to-one mapping of nonzero elements of $GF(9)$ on the indices $\pm 1, \pm 2, \pm 3, \pm 4$ is clearly arbitrary. Hence for no ordering of elements of $GF(9)$ on the circle can we define $\rho(\alpha, \beta)$ as the minimum number of arcs between α and β and have $\rho(\alpha, \beta)$ translation invariant. The above analysis applies with only trivial changes to arbitrary $GF(p^k)$ for p odd and $k = 2, 3, \cdots$.

Consider now the field $GF(2^k)$, $k \geq 3$. Let its 2^k distinct elements be denoted by

$$\{0 = \alpha_0, \alpha_i \; (i = \pm 1, \pm 2, \cdots, \pm(2^{k-1} - 1), 2^{k-1})\}.$$

Order these elements on a circle in the clockwise direction in the order 0, $\alpha_1, \alpha_2, \cdots, \alpha_{(2^{k-1}-1)}, \alpha_{2^{k-1}}, \alpha_{-(2^{k-1}-1)}, \cdots, \alpha_{-1}$, separating adjacent elements by unit arcs. Then define the Lee cyclic distance, $\rho(\alpha_i, \alpha_j)$ as the minimum number of arcs between α_i, α_j. Then $\rho(\alpha_i, \alpha_j)$ is not translation-invariant. Suppose the contrary. Then defining $|\alpha| = \rho(\alpha, 0)$, we have $\rho(\alpha, \beta) = |\alpha - \beta| = |\beta - \alpha|$, and $|\alpha|$ is a norm in the sense of Section II. Moreover, $|\alpha_i| = |i| =$ the absolute value of the integer i. Because $\alpha_i = -\alpha_i$, we have $\alpha_{-i} \neq -\alpha_i$ [which is different from the case $GF(p^k)$, p odd, treated above]. We have $|\alpha_2 - \alpha_1| = 1$. Hence $\alpha_2 - \alpha_1 = \alpha_1$ or α_{-1}. But $\alpha_2 - \alpha_1 = \alpha_1$ implies $\alpha_2 = 0$, which is a contradiction. Hence

$$\alpha_2 - \alpha_1 = \alpha_{-1}. \tag{9}$$

In addition, $|\alpha_3 - \alpha_2| = 1$ implies

$$\alpha_3 - \alpha_2 = \beta, \tag{10}$$

where $\beta = \alpha_1$ or $\beta = \alpha_{-1}$. Adding (9) and (10) we obtain

$$\alpha_3 - \alpha_1 = \alpha_{-1} + \beta.$$

If $\beta = \alpha_{-1}$, then $\alpha_3 = \alpha_1$, which is impossible. If $\beta = \alpha_1$, then $\alpha_3 = \alpha_{-1}$, which is impossible. Hence $\rho(\alpha, \beta)$ can not be translation-invariant. We summarize the above results as follows.

Theorem 11: Let $GF(q)$ be a Galois field of q elements where q is either 2^k for $k \geq 3$ or $q = p^k$, p odd and $k \geq 2$. Order the elements of $GF(q)$ on a circle with unit arc between pairs of adjacent elements, and define the distance, $\rho(\alpha, \beta)$, as the minimum number of arcs between α and β. Then $\rho(\alpha, \beta)$ is not translation-invariant in the sense of (4).

In order to extend the results of Section V to other nonprime fields, methods other than those of Section II are being investigated.

VII. Examples for the Case $GF(2^2)$

It seems worthwhile to give examples of the preceding discussion. Consider the field $GF(4)$ with the representation and laws of operation given in Section VI. Using the sieve method of Eratosthenes and trial and error it is easy to show that the only irreducible second degree polynomials with coefficients in $GF(4)$ which yield M-sequences are the following:

$$\omega x^2 + x + 1,$$
$$\omega^2 x^2 + x + 1,$$
$$\omega x^2 + \omega x + 1,$$
$$\omega^2 x^2 + \omega^2 x + 1.$$

These correspond respectively to the following recurren relations:

$$\alpha_k = \omega\alpha_{k-2} + \alpha_{k-1},$$
$$\alpha_k = \omega^2\alpha_{k-2} + \alpha_{k-1},$$
$$\alpha_k = \omega\alpha_{k-2} + \omega\alpha_{k-1},$$
$$\alpha_k = \omega^2\alpha_{k-2} + \omega^2\alpha_{k-1}.$$

As an example, we consider the relation $\alpha_k = \omega\alpha_{k-2} + \alpha_k$ which yields an M-sequence of period $4^2 - 1 = 15$ fro which we obtain the set (8) given by

$$a_0 = 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0,$$
$$a_1 = 1\ 1\ \omega^2\ 1\ 0\ \omega\ \omega\ 1\ \omega\ 0\ \omega^2\ \omega^2\ \omega\ \omega^2\ 0,$$
$$a_2 = 0\ 1\ 1\ \omega^2\ 1\ 0\ \omega\ \omega\ 1\ \omega\ 0\ \omega^2\ \omega^2\ \omega\ \omega^2$$

We can then show, for instance, that $a_1 - a_2 =$ and that

$$d(a_1, a_2) = 12 = (4 - 1)4^{2-1},$$
$$\delta(a_1, a_2) = 16 = 4^2,$$

where $d(a, b)$ is the Hamming distance, while $\delta(a, b)$ Lee's cyclic distance.

A Partial Ordering for Binary Channels*

N. ABRAMSON†, MEMBER, IRE

Summary—The properties of iterated binary channels are investi-ited. An ordering (defined by the symbol $\supset$) of communication iannels with two possible inputs and any number of possible itputs is defined. For any two such channels C_1 and C_2, this dering has the property that if $C_1 \supset C_2$, the minimum average ss when using C_1 will be less than the minimum average loss ıng C_2, *independent of the losses assigned to the various errors, ıd independent of the statistics of the source.* This ordering is plied to 1) the general binary channel, 2) the iterated binary mmetric channel, and 3) the unreliable binary symmetric channel hen used with many iterations. Curves allowing one to use the dering are given, and an example using these curves is worked ıt.

I. Comparison of Binary Channels

. *Introduction*

CONSIDER the use of a noisy binary channel to transmit information with a high degree of reliability. There are two common solutions to this roblem—block coding and iteration. In this paper we efine an ordering of binary channels pertinent when they re used with iteration.

Let a binary source emit symbols at the rate of one er second with the probability of a zero equal to ω, ıd the probability of a one equal to $(1 - \omega)$. Let us efine a channel C by the transition probability matrix[1]

$$C = \begin{bmatrix} q & 1-q \\ p & 1-p \end{bmatrix}, \tag{1}$$

ıd let C transmit at the rate of n binary symbols per cond. We may then transmit the message symbols om the source by iteration; that is, by transmitting oups of either n zeros or n ones through the channel C. Then used in this manner, we may think of the channel C generating another communication channel C_n. C_n is channel with two possible input messages—n zeros and ones—and 2^n possible output messages—the set of l n digit binary numbers. We shall refer to the channel $_n$ as the *nth-semi-extension* of C.

. *A Sufficient Statistic for C_n*

Let us call the two possible input messages to C_n, (n zeros), and s_2 (n ones). We also define the 2^n possible itput messages as $z_0, z_1, \cdots z_{2n-1}$ where z_j corresponds the n digit binary number j. Then it is possible to write a 2 by 2^n transition probability matrix for C_n,

$$\begin{bmatrix} P_{10} & P_{11} & \cdots & P_{1\,2^n-1} \\ P_{20} & P_{21} & \cdots & P_{2\,2^n-1} \end{bmatrix} \tag{2}$$

where P_{ij} is the conditional probability of z_j given s_i. P_{ij} is given by the equation

$$P_{ij} = \begin{cases} (1-q)^{w_j}(q)^{n-w_j} & i = 1 \\ (1-p)^{w_j}(p)^{n-w_j} & i = 2 \end{cases} \tag{3}$$

where

$$w_j = \text{number of ones in } z_j. \tag{4}$$

In order to decide which of the s_i was sent from an inspection of the received message, it is clearly not necessary to know *exactly* which of the z_j was received. It is necessary only to know w_j. That is, w_j is a sufficient statistic for the determination of the input message. We need, therefore, only distinguish $n + 1$ different outputs of the channel C_n — $w_j = 0, 1, \cdots n$. By considering the channel C_n to have two possible inputs—s_1 and s_2—and $n + 1$ possible outputs—0, 1, 2, $\cdots$ n—we may then write a simpler transition probability matrix as shown below.

$$\begin{bmatrix} (q)^n & \binom{n}{1}(q)^{n-1}(1-q) & \binom{n}{2}(q)^{n-2}(1-q)^2 & \cdots & (1-q)^n \\ (p)^n & \binom{n}{1}(p)^{n-1}(1-p) & \binom{n}{2}(p)^{n-2}(1-p)^2 & \cdots & (1-p)^n \end{bmatrix}. \tag{5}$$

C. *The Decision Theory Problem*

The problem of deciding which of the s_i was actually sent on the basis of which of the outputs $(0, 1 \cdots n)$ is received can be treated by the methods of statistical decision theory. We define a loss matrix

$$L = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} \tag{6}$$

where L_{ij} is the amount we lose if we decide that s_j was sent, if s_i was the true message sent. Then, remembering that the *a priori* probability of s_1 is ω, and the *a priori* probability of s_2 is $(1 - \omega)$, we may compute the optimum (Bayes) decision procedure—that procedure which minimizes the expected loss. Let us call the minimum value of the expected loss $R_n(\omega, L)$.

Now, if we are given the choice of using C_n, the nth-semi-extension of C or C'_m, the mth-semi-extension of some other channel C', it is a simple matter, in principle, to determine which channel to use. We need only compute $R_n(\omega, L)$ and $R'_m(\omega, L)$ and choose the channel with the smaller minimum expected loss. This method will generate a complete ordering of all possible semi-extensions of all

* Received by the PGIT, December 14, 1959. The work described this report was supported by the Office of Naval Res., Contract onr 225(24).

† Stanford Electronics Labs., Stanford, Calif.

[1] q is the probability of receiving a zero, if a zero is transmitted, d p is the probability of receiving a zero if a one is transmitted. e shall assume throughout the paper that $q \geq \frac{1}{2}$ and $p \leq \frac{1}{2}$.

possible binary channels. This ordering will, in general, *depend upon* ω *and* L. The surprising fact is that it is possible to assign a (partial) ordering to the C_n which is based upon minimum expected loss, yet which is *independent of* ω *and independent of* L. That is, in certain cases, it is possible to say that C_n is better than C_m' (that it will produce a smaller minimum average loss) *independent of the statistics of the source* (ω) *and independent of the losses we may assign to the different errors* (L).

If $R_n(\omega, L)$ is less than $R_m'(\omega, L)$ for all ω and all L, we say that C_n *is more informative than* C_m', *written* $C_n \supset C_m'$.

In the language of statistical decision theory, the determination of the transmitted message of C_n from the received message is a simple hypothesis testing problem. There are two *hypotheses*: H_1, or s_1 was sent, and H_2, or s_2 was sent. The observation of which of the outputs $(0, 1, 2, \cdots n)$ is received, is said to constitute an *experiment*, and the outputs $(0, 1, 2, \cdots n)$ are called the possible *outcomes of the experiment.*

It is important to note that by giving the transition probability matrix (5), we have completely defined the experiment corresponding to C_n. Henceforth, we shall refer to this matrix as the *experiment matrix*. The experiment matrix for the problem of testing hypotheses H_1 and H_2 is just a $2 \times k$ matrix (P_{ij}), where k is the number of possible outcomes of the experiment. P_{ij} is the conditional probability of outcome j, given that H_i is true. The problem of comparison of experiments—or equivalently comparison of experiment matrices—has been treated extensively in the statistical literature.[2-5] We shall follow the treatment of this subject as given by Blackwell.[2] A detailed explanation of this treatment is presented,[6] and a condensed version is given[7] in two articles by this author.

For the sake of completeness, we shall repeat[2] the condition for the comparison of two experiments in the next section.

D. Comparison of Dichotomous Experiments

Let P be an experiment with k possible outcomes $(x_1, x_2, \cdots x_k)$ used to test the hypotheses H_1 and H_2. P may be defined by the two by k matrix

$$\begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1k} \\ P_{21} & P_{22} & \cdots & P_{2k} \end{bmatrix} \tag{7}$$

where $P_{ij} = P_r\{x_j/H_i\}$; let

$$\alpha_j = P_{1j} + P_{2j}; \tag{8}$$

define

$$F_P(t) = \sum_{\left\{\alpha_j : \frac{P_{1j}}{\alpha_j} \le t\right\}} \alpha_j \tag{9}$$

(the notation on the right side of (9) indicates that fo[r] any t we sum only those α_j such that $P_{1j}/\alpha_j \le t$); finally define

$$K_P(t) = \int_0^t F_P(w)\, dw. \tag{10}$$

Then, if Q is some other experiment with the same numbe[r] of hypotheses, but not necessarily the same number o[f] outcomes,

$$P \supset Q \text{ if, and only if}$$
$$K_P(t) \ge K_Q(t) \text{ for all } t \text{ in } [0, 1]. \tag{11}$$

As suggested by (11), we shall call $K_P(t)$ the *compariso[n] function* of experiment P.

From the above definition, it may be seen that if $P \supset Q$ then $Q \subset P$, read Q is less informative than P. Furthermore, if

$$P \supset Q \quad \text{and} \quad Q \supset R$$

then

$$P \supset R.$$

That is, the relationship defined by $\supset$ is transitive. Finally, note that in general, given any two experiment[s] P and Q, we cannot say that either $P \supset Q$ or $Q \supset P$. When neither of these relations hold between experiment[s] P and Q, we say that P and Q are *noncomparable.* I[n] other words, the relationship defined by $\supset$ defines [a] partial ordering over all 2 by k Markov matrices,[8] o[r] equivalently over all channels with two possible transmitted messages and any number (not necessarily th[e] same for the channels being compared) of possible receive[d] messages.

II. Comparison of the General Binary Channel

A. The Comparison Function for C_1[9]

Let us apply the condition given in Section I-D t[o] two simple cases. First, consider the channel C_n whe[n] $n = 1$. That is, each signal, zero or one, is sent as it [is] received, and there is really no iteration (or codin[g]) at all. The experiment matrix for C_1 is just the transitio[n]

[2] D. Blackwell and M. Girshick, "Theory of Games and Statistical Decisions," John Wiley and Sons, Inc., New York, N. Y.; 1954.

[3] D. Blackwell, "Equivalent comparisons of experiments," *Ann. Math. Stat.*, vol. 24, pp. 265–272; 1953.

[4] D. Lindley, "On a measure of the information provided by an experiment," *Ann. Math. Stat.*, vol. 27, pp. 986–1005; 1956.

[5] R. Bradt and S. Karlin, "On the design and comparison of certain dichotomous experiments," *Ann. Math. Stat.*, vol. 27, pp. 390–409; 1956.

[6] N. Abramson, "Application of 'Comparison of Experiments' to Radar Detection and Coding Problems," Stanford Electronics Labs., Stanford, Calif., Tech. Rept. No. 41; July 28, 1958.

[7] N. Abramson, "The application of comparison of experiments to detection problems," 1958 IRE National Convention Record, pt. 4, pp. 22–26.

[8] This ordering may also be extended to include more gener[al] experiments. See footnote 3.

[9] We shall adopt the notation

$$F_n(t) \text{ for } F_{C_n}(t)$$

and

$$K_n(t) \text{ for } K_{C_n}(t).$$

obability matrix

$$\begin{bmatrix} q & 1-q \\ p & 1-p \end{bmatrix}. \tag{12}$$

From (8), we have[10]

$$\alpha_1 = q + p$$

$$\alpha_2 = (1-q) + (1-p),$$

d from (9), remembering that $q \geq \frac{1}{2}$ and $p \leq \frac{1}{2}$,

$$F_1(t) = \begin{cases} 0 & t < t^{(1)} \\ (1-q) + (1-p) & t^{(1)} \leq t < t^{(2)} \\ 2 & t^{(2)} \leq t \end{cases} \tag{13}$$

ıere

$$t^{(1)} \triangleq \frac{1-q}{(1-q)+(1-p)} \tag{14a}$$

$$t^{(2)} \triangleq \frac{q}{q+p}. \tag{14b}$$

nally, we may plot $K_1(t)$, the comparison function of , in Fig. 1.

The Comparison of Two Binary Channels

From (11) we see that $C_1 \supset C_1'$ for any two noniterated ıary channels C_1 and C_1' if, and only if

$$K_1(t) \geq K_1'(t) \quad \text{for all } t \text{ in } \quad [0, 1]; \tag{15}$$

d after a little algebra, we see that a necessary and fficient condition for (15) to hold is

$$t^{(1)} \leq t^{(1)\prime} \tag{16a}$$

d

$$t^{(2)} \geq t^{(2)\prime}. \tag{16b}$$

nally, using (14), we may express (16) as

$$\frac{p}{q} \leq \frac{p'}{q'} \tag{17a}$$

d

$$\frac{1-p}{1-q} \geq \frac{1-p'}{1-q'}. \tag{17b}$$

Eq. (17) may be expressed graphically as in Fig. 2. Any channel C_1 with experiment matrix as in (12) ıy be thought of as a point in the unit square. Strictly eaking, since we have assumed that $q \geq \frac{1}{2}$ and $p \leq \frac{1}{2}$ should only consider points in the upper left-hand arter of the unit square. In Fig. 2 we have fixed the int (p', q') corresponding to the channel C_1'. The set channels whose parameters p and q satisfy both (17a) d (17b) will then lie in the shaded region of Fig. 2.

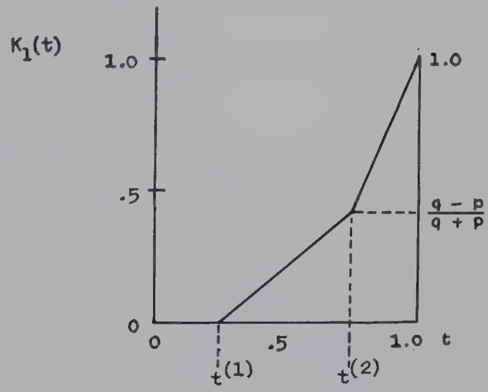


Fig. 1—Comparison function for the general binary channel.

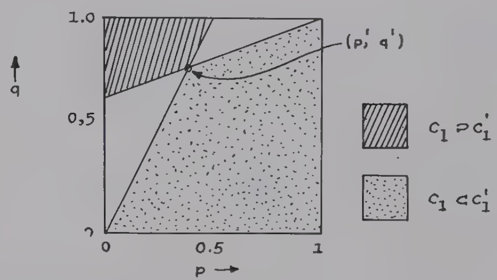


Fig. 2—Comparison of C_1 and C_1'.

These channels are all the channels which are *more* informative than C_1'. The set of channels whose parameters p and q satisfy (17a) and (17b) with the inequality signs reversed will lie in the dotted region of Fig. 2. These channels are all the channels which are less informative than C_1'. The channels corresponding to the unshaded and undotted regions of Fig. 2 are those channels which are noncomparable with C_1'.

C. The Shannon Ordering

For the case of the noniterated binary channel, it is interesting to compare the partial ordering presented in the previous section with a different partial ordering discussed by Shannon.[11] We shall use the symbol $\succ$ to denote Shannon's ordering. It is easily shown that if we have two channels C_1 and C_1', then $C_1 \supset C_1'$ implies $C_1 \succ C_1'$. The converse, however, is not true. For the case of the binary *symmetric* channel, it may be seen that the two orderings are equivalent.

III. Comparison of the Iterated BSC

A. The Comparison Function for the Iterated BSC

Now consider the channel C_n, the nth semi-extension of C_1. The experiment matrix for C_n is given in (5). In principle we may start with the experiment matrix and calculate its comparison function $K_n(t)$. To compare C_n

[10] This simple example, until (16), is taken with some notational ınges, from Blackwell (see footnote 2). It is repeated here for lagogical reasons.

[11] C. Shannon, "A note on a partial ordering for communication channels," *Information and Control*, vol. 1, pp. 390–397; December, 1958.

with C_m' the mth semi-extension of C_1' we need only examine

$$K_n(t) - K_m'(t)$$

and note whether this difference is non-negative or non-positive in the interval [0, 1]. In practice, however, for large n such a procedure can become quite tedious. For the case of the nth semi-extension of a binary symmetric channel (BSC), however, there does exist an easy method of obtaining the comparison function. For the BSC the transition probability matrix is

$$C = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \tag{18}$$

and the experiment matrix of the C_n, the nth semi-extension of this channel is just.

$$\begin{bmatrix} (1-p)^n & \binom{n}{1}(1-p)^{n-1}p & \binom{n}{2}(1-p)^{n-2}(p)^2 & \cdots & (p)^n \\ (p)^n & \binom{n}{1}(p)^{n-1}(1-p) & \binom{n}{2}(p)^{n-2}(1-p)^2 & \cdots & (1-p)^n \end{bmatrix}. \tag{19}$$

Now, using the notation of Section I-D we see that

$$\alpha_j = \binom{n}{j-1}[(1-p)^{n-j+1}(p)^{j-1} + (1-p)^{j-1}(p)^{n-j+1}], \qquad j = (1, 2, \cdots n+1) \tag{20}$$

and defining

$$r \triangleq \frac{p}{1-p} \leq 1 \tag{21}$$

we may write (20) as

$$\alpha_j = \binom{n}{j-1}(1-p)^n(r^{j-1} + r^{n-j+1}) \qquad j = (1, 2, \cdots n+1). \tag{22}$$

Further, as suggested by (9), we define

$$t_j \triangleq \frac{\binom{n}{j-1}(1-p)^{n-j+1}(p)^{j-1}}{\alpha_j} \qquad j = (1, 2, \cdots n+1) \tag{23a}$$

and, simplifying (23a), we obtain

$$t_j = \frac{1}{1 + r^{n-2j+2}} \qquad j = (1, 2, \cdots n+1). \tag{23b}$$

Now the subscripts on t_j and α_j in (22) and (23) refer to the jth column of the experiment matrix. It will be convenient to re-lable the t_j and α_j as follows:

$$\begin{array}{ll} t_1 \to t^{(n+1)} & \alpha_1 \to \alpha^{(n+1)} \\ t_2 \to t^{(n)} & \alpha_2 \to \alpha^{(n)} \\ \vdots & \vdots \\ t_{n+1} \to t^{(1)} & \alpha_{n+1} \to \alpha^{(1)} \end{array} \tag{26}$$

so that (23b) may be rewritten as

$$t^{(j)} = \frac{1}{1 + r^{2j-(n+2)}}. \qquad (23$$

Note that now (since $r \leq 1$)

$$0 \leq t^{(1)} \leq t^{(2)} \leq \cdots \leq t^{(n+1)} \leq 1, \qquad (2$$

so that from (9) we may write

$$\left(\text{remember that} \quad \sum_{j=1}^{n+1} \alpha^{(j)} = 2\right)$$

$$F_n(t) = \begin{cases} 0 & 0 \leq t < t^{(1)} \\ \alpha^{(1)} & t^{(1)} \leq t < t^{(2)} \\ \alpha^{(1)} + \alpha^{(2)} & t^{(2)} \leq t < t^{(3)} \\ \alpha^{(1)} + \alpha^{(2)} + \alpha^{(3)} & t^{(3)} \leq t < t^{(4)} \\ \vdots & \vdots \\ 2 - \alpha^{(n+1)} & t^{(n)} \leq t < t^{(n+1)} \\ 2 & t^{(n+1)} \leq t \leq 1 \end{cases}. \qquad (2$$

And the comparison function may be written as

$$K_n(t) = \int_0^t F_n(w)\, dw = \sum_{k=1}^{j} \alpha^{(k)}[t - t^{(k)}] \qquad (2$$

$$\text{for} \quad t^{(j)} \leq t \leq t^{(j+1)}$$

$$\text{and} \quad j = 0, 1, 2, \cdots n+1$$

where we define $t^{(0)} = 0$ and $t^{(n+2)} = 1$.

Note that $K_n(t)$ is also a function of r. When we wi to emphasize this dependence upon r we shall wr $K_n(t)$ as $K_{n,r}(t)$. $t^{(i)}$ is a function of n and r. When wish to emphasize the dependence upon n we shall wr $t^{(i)}$ as $t_{n(i)}$. When we wish to emphasize the dependen upon both n and r, we shall write $t_{n,r}^{(i)}$. The first su script will always refer to the number of iterations, t second to the quantity $p/(1-p)$.

B. Four Lemmas

Eq. (29) provides one possible method of computi the comparison function for the nth semi-extension a BSC. In this section we shall state several lemm which may be used to decrease considerably the amou of work involved in this computation. We leave the pro of these lemmas to the appendix.

Lemma 1: Let $K_{n,r}(t)$ and $K_{m,r'}(t)$ be the compari functions of the nth and mth semi-extension of the BS $C_{1,r}$ and $C_{1,r'}$ respectively.

If

$$K_{n,r}(t) \leq K_{m,r'}(t) \quad \text{for} \quad t \text{ in } [0, \tfrac{1}{2}]$$

en

$$K_{n,r}(t) \leq K_{m,r'}(t) \quad \text{for} \quad t \text{ in } [0, 1].$$

Lemma 1 states that in order to compare iterations of 3C's, we need only compute the comparison functions r $0 \leq t \leq \frac{1}{2}$.

Next note the form of $K_n(t)$ as given in (29). $K_n(t)$ a piecewise linear function of t in the interval [0, 1]. ırthermore $K_n(t)$ is continuous in this interval, while (t), the derivative of $K_n(t)$, is discontinuous at the ints $t^{(1)}$, $t^{(2)}$, $\cdots$ $t^{(n+1)}$. We shall refer to these points the "breakpoints" of $K_n(t)$.

The breakpoints of $K_n(t)$ are easily obtained from (23c). addition by inspection of (23c) we may immediately ite Lemma 2.

Lemma 2: Let $K_n(t)$ and $K_{n-2}(t)$ be the comparison ıctions of the nth and $n - 2$ th semi-extensions of *same* BSC. The set of $n + 1$ breakpoints of $K_n(t)$ nsist of

a) the set of $n - 1$ breakpoints of $K_{n-2}(t)$,

b) $$\frac{1}{1 + r^{-n}},$$

c) $$\frac{1}{1 + r^{n}} = 1 - \frac{1}{1 + r^{-n}}.$$

Lemma 2 states that if we have the $n - 1$ breakpoints $K_{n-2}(t)$, it is only necessary to compute one additional eakpoint to obtain the $n + 1$ breakpoints of $K_n(t)$. mma 2 suggests that a simple method of obtaining a nparison function for large n might be to start with aller values of n and to proceed by induction. This leed will be the method we shall use but first we need o additional results.

Lemma 3: If $K_n(t)$ is the comparison function for the ı semi-extension of a BSC then, for n an odd integer

$$\left[\frac{d}{dt} K_n(t)\right]_{t=1/2} = 1.$$

The last result we need before we show how to obtain comparison function for the nth semi-extension of 3SC is Lemma 4.

Lemma 4: Let $K_n(t)$ and $K_{n-1}(t)$ be the comparison ıctions of the nth and $n - 1$th semi-extensions of *same* BSC. Let $t_n^{(1)}$, $t_n^{(2)}$, $\cdots$ $t_n^{(n+1)}$ be the $n + 1$ breakints of $K_n(t)$. Then

$$K_n(t_n^{(j)}) = K_{n-1}(t_n^{(j)}) \quad \text{for} \quad n = 2, 3, \cdots \quad \text{and} \quad j = 1, 2, \cdots n + 1.$$

Lemma 4 tells us that $K_n(t)$ will equal $K_{n-1}(t)$ at the eakpoints of $K_n(t)$. Thus, if we know $K_{n-1}(t)$ we know value of $K_n(t)$ at its breakpoints. $K_n(t)$ is, however, inear function of t between any two successive breakpoints so that the values of $K_n(t)$ at its breakpoints are sufficient to determine the function everywhere.

C. *Construction of the Comparison Function for an Iterated BSC*

We shall now illustrate the application of Lemmas 1, 2, 3 and 4 to the construction of a comparison function. As an example, let us say we wish to construct $K_n(t)$ for $n = 10$ and $r = 0.5$. Note that by Lemma 1, we are only interested in $K_n(t)$ for $0 \leq t \leq \frac{1}{2}$. Now the breakpoints of $K_{10}(t)$ and $K_9(t)$ may be computed from (23c). In Table I we have listed for these two comparison functions all the breakpoints which occur in the interval $[0, \frac{1}{2}]$.

TABLE I
BREAKPOINTS OF $K_{10}(t)$ AND $K_9(t)$ FOR $r = 0.5$

	$K_{10}(t)$	$K_9(t)$
$t^{(1)}$	0.001	0.002
$t^{(2)}$	0.004	0.008
$t^{(3)}$	0.150	0.300
$t^{(4)}$	0.590	0.111
$t^{(5)}$	0.200	0.333
$t^{(6)}$	0.500	

By Lemma 2 the breakpoints of $K_8(t)$ occurring in the interval $[0, \frac{1}{2}]$ may be obtained by dropping the first entry in the first column of Table I. The breakpoints of $K_6(t)$ may be obtained by dropping the first two entries in the first column, etc. The breakpoints for $K_7(t)$, $K_5(t)$, $K_3(t)$ and $K_1(t)$ are obtained from the second column in Table I in the same manner.

The first step in the construction of $K_{10}(t)$ then is to indicate the breakpoints of $K_1(t)$, $K_2(t)$, $\cdots$ $K_{10}(t)$ in the interval $[0, \frac{1}{2}]$. These breakpoints, which by Lemma 2 consist of the eleven numbers in Table I, are shown by vertical lines in Fig. 3.

Having drawn these lines, it is a simple matter to construct $K_1(t)$ for t in $[0, \frac{1}{2}]$. The only breakpoint of $K_1(t)$ in $[0, \frac{1}{2}]$ is the last entry in the second column of Table I ($t = 0.333$). By Lemma 3, $K_1(t)$ must have a slope of unity at $t = \frac{1}{2}$. We therefore draw the line starting at this breakpoint with unit slope. This line is equal to $K_1(t)$ for $0.333 \leq t \leq 0.500$. $K_1(t)$ is equal to zero for $0 \leq t \leq 0.333$. $K_2(t)$ may now be drawn immediately as shown in Fig. 3 by a direct application of Lemma 4. In exactly the same manner, by using Lemma 4, we construct $K_3(t)$ to $K_{10}(t)$ without further calculation.[12]

In Figs. 4–6 we have drawn the comparison functions for some other values of r.

D. *Example*

As an example of the application of these figures, consider the use of the third extension of a BSC with proba-

[12] It is necessary to employ the slope condition given in Lemma 3 to obtain the last segment of $K_j(t)$ for odd j.

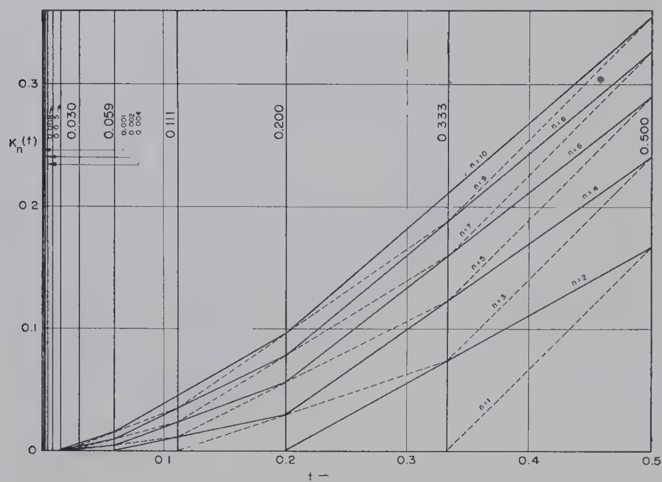


Fig. 3—Comparison functions, $K_n(t)$, for $r = 0.5$.

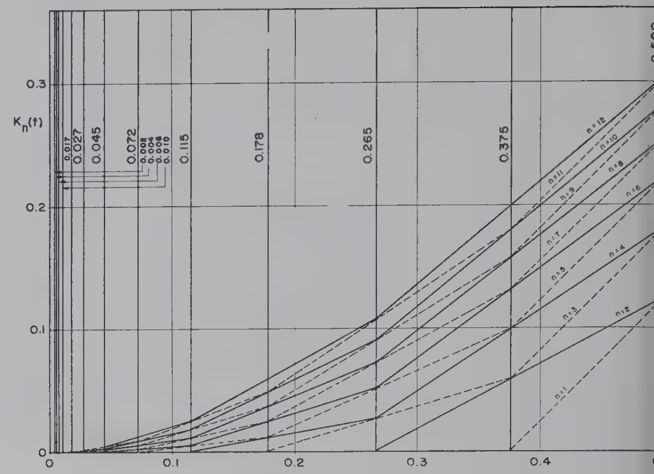


Fig. 4—Comparison functions, $K_n(t)$, for $r = 0.6$.

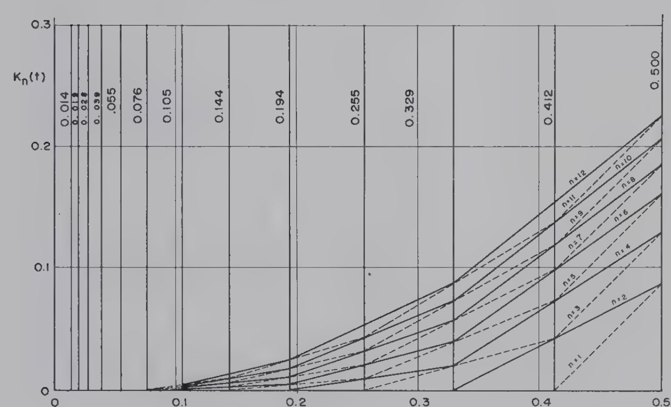


Fig. 5—Comparison functions, $K_n(t)$, for $r = 0.7$.

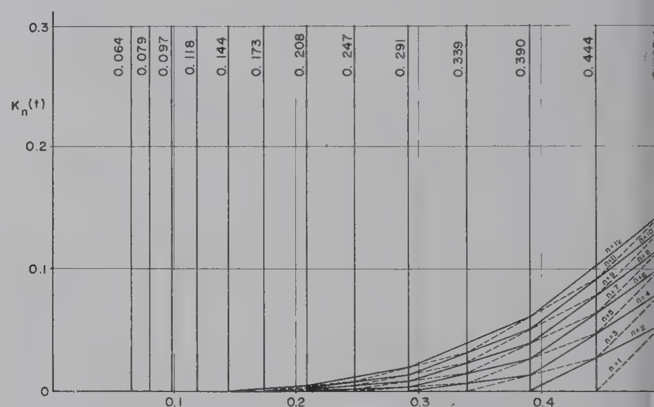


Fig. 6—Comparison functions, $K_n(t)$, for $r = 0.8$.

bility of error equal to 0.375 ($r = 0.6$). By superimposing Fig. 4 on Fig. 5 it may be seen that

$$K_{3,.6}(t) \geq K_{4,.7}(t) \qquad 0 \leq t \leq \tfrac{1}{2}$$

and

$$K_{3,.6}(t) \leq K_{8,.7}(t) \qquad 0 \leq t \leq \tfrac{1}{2}.$$

That is, if the probability of error of the BSC increases to 0.411 ($r = 0.7$) the first channel when used with three repetitions will *always* be preferred to the second channel even if we use it with as many as four repetitions. On the other hand the first channel when used with three repetitions will *never* be preferred to the second channel when used with eight or more repetitions.

E. Behavior For Large N

By inspection of the comparison functions in Figs. 3–6 (or from Lemma 4) it may be seen that for a fixed r, $K_{n+1}(t) \geq K_n(t)$. That is, the $n + 1$th semi-extension of a BSC is more informative than the nth semi-extension of the same BSC. This not too surprising fact brings up the question of the behavior of $K_n(t)$ for large n. Does $K_n(t)$ approach a limit and if so is this limit a function of r? These questions are answered by Theorem 1.

Theorem 1.[13] Let $K_n(t)$ be the comparison functi of the nth semi-extension of a BSC with transition pro bility matrix

$$\begin{bmatrix} 1 - p & p \\ p & 1 - p \end{bmatrix}.$$

a) If $p < \frac{1}{2}$

$$\lim_{n\to\infty} K_n(t) = t \qquad 0 \leq t \leq 1;$$

b) if $p = \frac{1}{2}$

$$\lim_{n\to\infty} K\ (t) = \begin{cases} 0 & 0 \leq t \leq \frac{1}{2} \\ 2t - 1 & \frac{1}{2} \leq t \leq 1 \end{cases}.$$

We have plotted these two limits in Fig. 7. We s call the greater of these limits the *maximum compar function* (MXCF) and the other the *minimum compar function* (MNCF). It is not possible for a channel w two inputs and any number of outputs to have a c parison function *greater* than the MXCF at *any* p in [0, 1]. Likewise, it is not possible for a channel w two inputs and any number of outputs to have a c

[13] Theorem 1 is proved in the appendix.

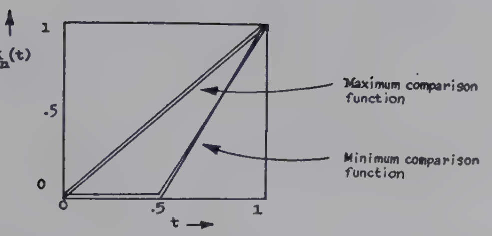


Fig. 7—The maximum and minimum comparison functions.

arison function *less* than the MNCF at *any* point in [0, 1]. he MXCF corresponds to a channel which can transmit inary information with zero error. The MNCF corresponds to a channel where the output is statistically ndependent of the input.

IV. Comparison of the Iterated BSC With C_1'

. Two Comparison Theorems

In this section we shall investigate the comparison of C_n, he nth semi-extension of one BSC, with C_1', some other *oniterated* BSC. By means of such a comparison it will e possible to compare iterated BSC's in general. That is

$$\text{if} \quad C_n \supset C_1'$$
$$\text{and} \quad C_1' \supset C_m'',$$

hen we know that

$$C_n \supset C_m''.$$

The reason for comparing two iterated BSC's through e artifice of introducing another, noniterated BSC may e seen in the following lemma.[14]

Lemma 5: Let C_n be the nth semi-extension of the SC C_1 and let C_1' be some other BSC. Then

a) $C_1' \supset C_n$ if and only if

$$t_{1,r'}^{(1)} \le t_{n,r}^{(1)}. \tag{30}$$

b) $C_n \supset C_1'$ if and only if

$$K_{n,r}(\tfrac{1}{2}) \ge K_{1,r'}(\tfrac{1}{2}). \tag{31}$$

We are now in position to state the two central results Section III. First, by using in (30) the expression for $^{j)}$ given in (23c), and then writing the result in terms f p, we obtain Theorem 2, as follows:

Theorem 2: Let C_1 and C_1' be two BSC's with probailities of error p and p' respectively. Let C_n be the nth mi-extension of C_1. Then

$C_1' \supset C_n$ if and only if

$$p' \le \frac{(p)^n}{(p)^n + (1-p)^n}. \tag{32}$$

Next we note that for n even $t_n^{n/2+1} = \frac{1}{2}$, so that y Lemma 4, $K_n(\frac{1}{2}) = K_{n-1}(\frac{1}{2})$, again for even n. Then get $K_{n-1}(\frac{1}{2})$, we use (29), setting $t = \frac{1}{2}$, and finally btain:

[14] Lemma 5 is proved in the appendix.

Theorem 3: Let C_1, and C_1' be two BSC's with probabilities of error p and p' respectively. Let C_n be the nth semi-extension of C_1. Then

a) for n odd

$C_n \supset C_1'$ if and only if

$$p' \ge \sum_{k=1}^{(n+1)/2} \binom{n}{k-1} (p)^{n-k+1}(1-p)^{k-1}, \tag{33a}$$

b) for n even

$C_n \supset C_1'$ if and only if

$$C_{n-1} \supset C_1'. \tag{33b}$$

B. Some Comparison Regions

In most of the remainder of Section IV we shall be interested in investigating the properties of highly unreliable BSC's when used with a large number of iterations. Accordingly, we shall find it convenient to define the *BSC parameter* ϵ by

$$p = \tfrac{1}{2}(1-\epsilon) \tag{34}$$

where p is the probability of error of the BSC. We recall that

$$r = \frac{p}{1-p} \tag{21}$$

so that

$$r = \frac{1-\epsilon}{1+\epsilon}. \tag{35}$$

For a highly unreliable BSC the BSC parameter ϵ will approach zero.

Let ϵ and ϵ' be the BSC parameters of C_1 and C_1' respectively. Then (32) may be written as

$C_1' \supset C_n$ if and only if

$$\frac{1-\epsilon'}{2} \le \frac{(1-\epsilon)^n}{(1-\epsilon)^n + (1+\epsilon)^n} \tag{36a}$$

or

$C_1' \supset C_n$ if and only if

$$\epsilon' \ge \frac{\binom{n}{1}\epsilon + \binom{n}{3}\epsilon^3 + \binom{n}{5}\epsilon^5 + \cdots}{1 + \binom{n}{2}\epsilon^2 + \binom{n}{4}\epsilon^4 + \cdots} \tag{36b}$$

where we have defined

$$\binom{n}{m} = 0 \quad \text{for} \quad m > n.$$

Likewise, we may rewrite (33a) in terms of ϵ and ϵ' as follows: For n odd

$C_n \supset C_1'$ if and only if

$$\epsilon' \le 1 - \left(\frac{1}{2}\right)^{n-1} \sum_{k=1}^{(n+1)/2} \binom{n}{k-1} (1+\epsilon)^{k-1}(1-\epsilon)^{n-k+1}. \tag{37}$$

In Figs. 8–11 we have plotted in the ϵ, ϵ' plane regions where $C_1' \supset C_n$ and where $C_n \supset C_1'$ for $n = 3, 5, 7$ and 9. Also included in these figures is the *equicapacity line*—the line which gives for any ϵ' the value of ϵ such that the Shannon channel capacity of C_n (in bits per unit time) is equal to that of C_1'. The channel capacity of C_n is given by

$$n[1 + p \log p + (1 - p) \log (1 - p)] \qquad (38)$$

where we have multiplied by n so that one digit from the source may be transmitted by C_n (using n iterations) in the same time that it may be transmitted by C_1'. The channel capacity of C_1' is, of course,

$$1 + p' \log p' + (1 - p') \log (1 - p'). \qquad (3$$

It is interesting to note that the equicapacity li lies in the noncomparable regions of Figs. 8–11 exce for a small range of values of ϵ' corresponding to high reliable BSC's. This range of values vanishes rapidly n increases. In this range, however, we have the i teresting phenomenon of two channels C_n and C', wi the channel capacity of C_n greater than that of C' ar yet with the average loss using C_n (in an iterated manne always greater than that of C'.

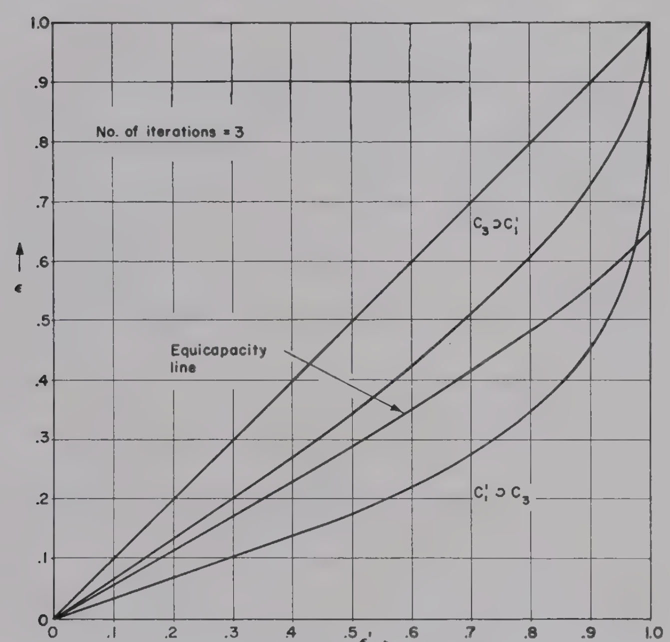


Fig. 8—Comparison of the iterated BSC with the noniterated BSC.

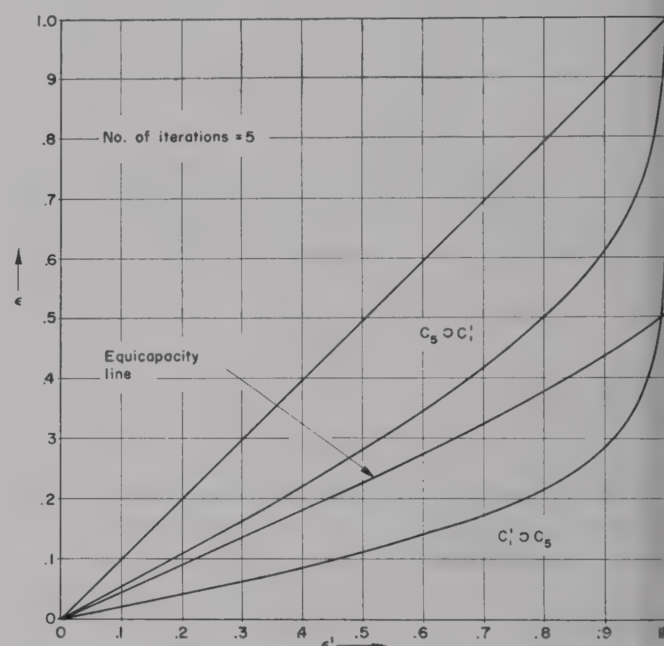


Fig. 9—Comparison of the iterated BSC with the noniterated BS

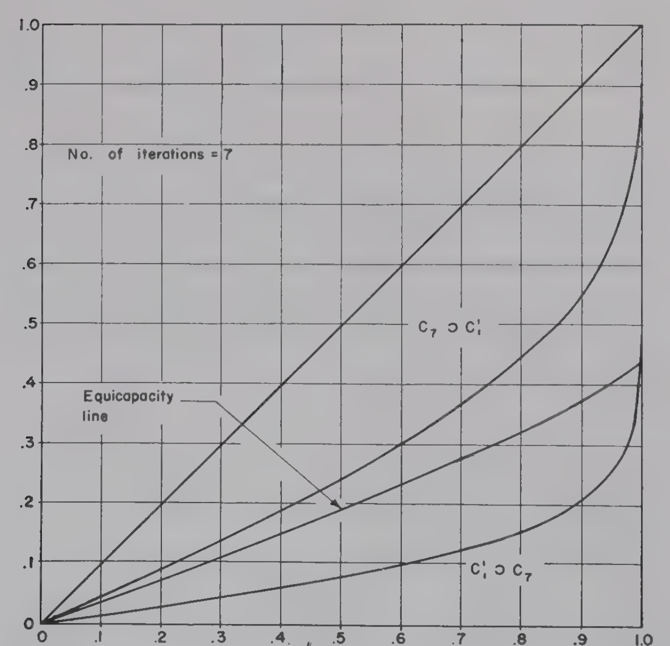


Fig. 10—Comparison of the iterated BSC with the noniterated BSC.

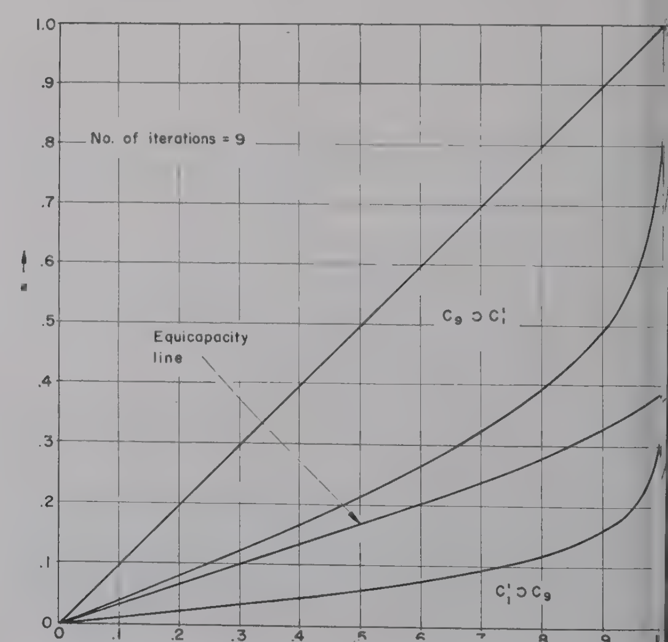


Fig. 11—Comparison of the iterated BSC with the noniterated B

. Iteration of Unreliable BSC's

One final question of interest is the behavior of highly nreliable channels when used with many iterations. That , we are interested in comparing C_n with C_1' when ϵ pproaches zero and n is a large number.

Under these conditions (36b) reduces to

$$C_1' \supset C_n \text{ if and only if}$$

$$\epsilon \leq \frac{\epsilon'}{n}. \tag{40}$$

We may also let ϵ' approach zero in (37). After a good eal of algebraic manipulation and the application of tirling's formula, this will reduce to

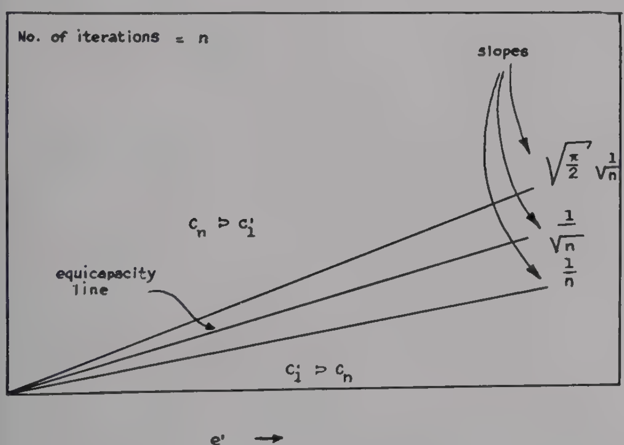
$$C_n \supset C_1' \text{ if and only if}$$

$$\epsilon \geq \sqrt{\frac{\pi}{2}}\frac{\epsilon'}{\sqrt{n}}. \tag{41}$$

Finally from (38) and (39) we obtain (again for small ϵ): The capacity of C_n [from (38)] is equal to the capacity f C_1' [from (39)] when

$$\epsilon = \frac{\epsilon'}{\sqrt{n}}. \tag{42}$$

We may indicate the results of (40)–(42) as shown in ig. 12.



g. 12—Comparison of the iterated unreliable BSC with the noniterated BSC.

Appendix

ome Proofs

Lemma 1:

Let $K_{n,r}(t)$ and $K_{m,r'}(t)$ be the comparison functions the nth and mth semi-extensions of the BSC's $C_{1,r}$ nd $C_{1,r'}$ respectively.

If

$$K_{n,r}(t) \geq K_{m,r'}(t) \quad \text{for} \quad t \text{ in } [0, \tfrac{1}{2}]$$

hen

$$K_{n,r}(t) \geq K_{m,r'}(t) \quad \text{for} \quad t \text{ in } [0, 1].$$

Proof: From (10)

$$K_n(t) = \int_0^t F_n(w)\,dw \tag{43}$$

and letting

$$v = 1 - w$$

$$K_n(t) = \int_{1-t}^1 F_n(1 - v)\,dv, \tag{44}$$

but from (28) we see that if $K_n(t)$ is the comparison function of the nth semi-extension of a BSC, then

$$F_n(1 - v) = 2 - F_n(v) \quad \text{for} \quad v \text{ in } [0, 1] \tag{45}$$

except on a finite set of points, so that

$$\begin{aligned} K_n(t) &= \int_{1-t}^1 [2 - F_n(v)]\,dv \\ &= 2t - K_n(1) + K_n(1 - t) \\ &= (2t - 1) + K_n(1 - t). \end{aligned} \tag{46}$$

Lemma 1 then follows directly from (46).

Lemma 2:

Let $K_n(t)$ and $K_{n-2}(t)$ be the comparison functions of the nth and $n - 2$th semi-extensions of the *same* BSC. The set of $n + 1$ breakpoints of $K_n(t)$ consist of

a) the set of $n - 1$ breakpoints of $K_{n-2}(t)$

b) $$\frac{1}{1 + r^{-n}}$$

c) $$\frac{1}{1 + r^n} = 1 - \frac{1}{1 + r^{-n}}.$$

Proof: Lemma 2 follows directly from (23c).

Lemma 3:

If $K_n(t)$ is the comparison function for the nth semi-extension of a BSC then, for n an odd integer

$$\left[\frac{d}{dt}K_n(t)\right]_{t=1/2} = 1.$$

Proof: For n odd we may use (29) where $j = (n + 1)/2$. Then, taking a derivative with respect to t, we get

$$\left[\frac{d}{dt}K_n(t)\right]_{t=1/2} = \sum_{k=1}^{(n+1)/2} \alpha^{(k)}. \tag{47}$$

Finally, from (19) we see that the right side of (47) is just one.

Lemma 4:

Let $K_n(t)$ and $K_{n-1}(t)$ be the comparison functions of the nth and $n - 1$th semi-extensions of the *same* BSC. Let $t_n^{(1)}, t_n^{(2)}, \cdots t_n^{(n+1)}$ be the $n + 1$ breakpoints of $K_n(t)$.

Then

$$K_n(t_n^{(j)}) = K_{n-1}(t_n^{(j)})$$
$$\text{for } n = 2, 3, \cdots$$
$$\text{and } j = 1, 2, \cdots n+1.$$

Proof: The first step is to note that

$$K_n(t_n^{(1)}) = K_{n-1}(t_n^{(1)}) = 0; \tag{48}$$

next we show that

$$K_n(t_n^{(2)}) = K_{n-1}(t_n^{(2)}). \tag{49}$$

$K_n(t)$ is just the integral of $F_n(t)$. In Fig. 13 we have plotted the first section of $F_n(t)$ and $F_{n-1}(t)$. Referring to this figure we see that $K_n(t_n^{(2)})$ is just the area under $F_n(t)$ from 0 to $t_n^{(2)}$ while $K_{n-1}(t_n^{(2)})$ is the area under $F_{n-1}(t)$ in this same interval. Thus to show that these two areas are equal we need only show that the shaded area A_1 in Fig. 13 is equal to the shaded area B_1. That is, we must prove that

$$[t_{n-1}^{(1)} - t_n^{(1)}][\alpha_n^{(1)}] = [t_n^{(2)} - t_{n-1}^{(1)}][\alpha_{n-1}^{(1)} - \alpha_n^{(1)}]. \tag{50}$$

Substituting for the $t_i^{(k)}$ and $\alpha_i^{(k)}$ in this equation will then prove (49).

We might continue in this manner and show that areas A_2 and B_2 of Fig. 13 are equal so that $K_n(t)$ and $K_{n-1}(t)$ are equal for $t = t_n^{(3)}$. It is easier, however, to tackle the general problem immediately. That is, referring to Fig. 14, we shall show that the fact that area A_{j-1} is equal to area B_{j-1} implies that area A_j is equal to area B_j.

In terms of the quantities in Fig. 14 we wish to show that if

$$[L_{j-1}][t_n^{(j)} - t_{n-1}^{(j-1)}] = [\alpha_{n-1}^{(j-1)} - L_{j-1}][t_{n-1}^{(j-1)} - t_n^{(j-1)}] \tag{51}$$

then we must have

$$[t_n^{(j+1)} - t_{n-1}^{(j)}][\alpha_{n-1}^{(j)} - \alpha_n^{(j)} + L_{j-1}] = [t_{n-1}^{(j)} - t_n^{(j)}][\alpha_n^{(j)} - L_{j-1}]. \tag{52}$$

From (51) we may obtain an expression for L_{j-1}

$$L_{j-1} = \alpha_{n-1}^{(j-1)} \frac{t_{n-1}^{(j-1)} - t_n^{(j-1)}}{t_n^{(j)} - t_n^{(j-1)}}; \tag{53}$$

after substitution of (53) in (52) and a good deal of algebra we obtain

$$\alpha_n^{(j)} \stackrel{?}{=} \alpha_{n-1}^{(j)} \frac{t_n^{(j+1)} - t_{n-1}^{(j)}}{t_n^{(j+1)} - t_n^{(j)}} + \alpha_{n-1}^{(j-1)} \frac{t_{n-1}^{(j-1)} - t_n^{(j-1)}}{t_n^{(j)} - t_n^{(j-1)}}; \tag{54}$$

finally, we substitute for $t_i^{(k)}$ and $\alpha_i^{(k)}$ and verify (54).

Let us summarize what we have done. First we have shown that

$$K_n(t_n^{(j)}) = K_{n-1}(t_n^{(j)}) \tag{55}$$

for $j = 1$ and 2. Eq. (55) would hold for all j if we could prove that the areas we have labeled A_j and B_j were equal. Finally we proved that $A_j = B_j$ if $A_{j-1} = B_{j-1}$; since we had already proved $A_1 = B_1$ this completed the proof and (55) holds for all j.

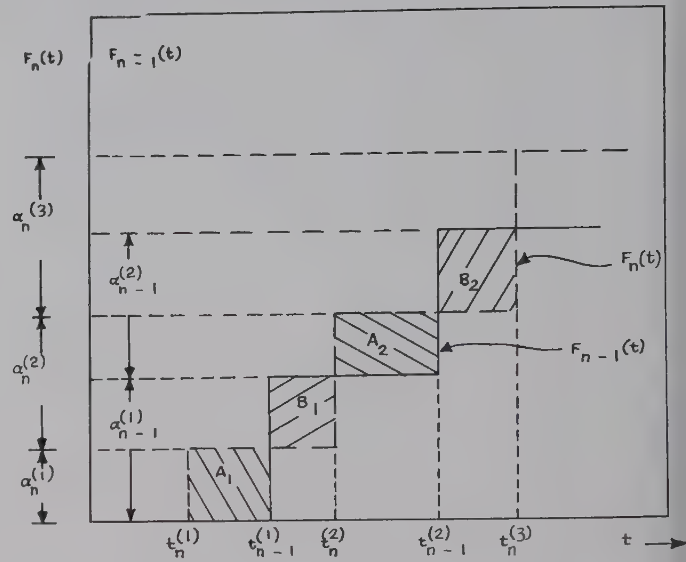


Fig. 13—The first section of $F_n(t)$ and $F_{n-1}(t)$.

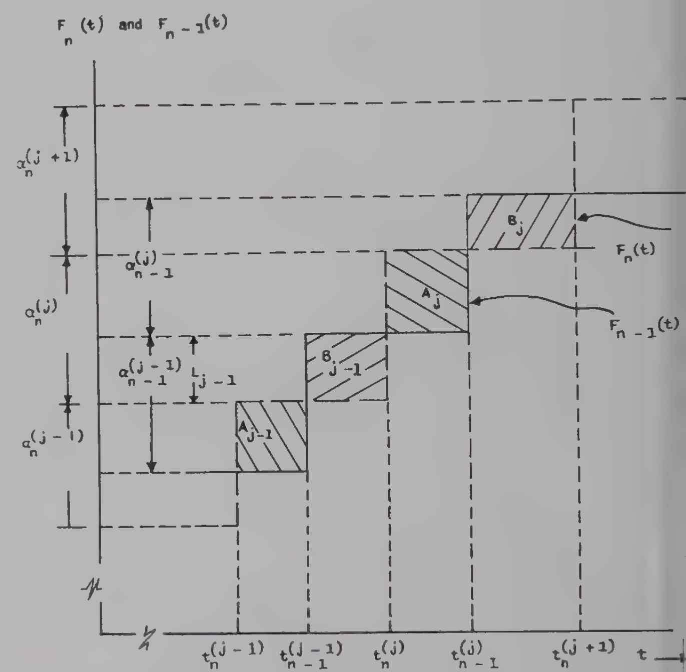


Fig. 14—A middle section of $F_n(t)$ and $F_{n-1}(t)$.

Theorem 1:

Let $K_n(t)$ be the comparison function of the nth semi extension of a BSC with transition probability matrix

$$\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

a) if $p < \frac{1}{2}$

$$\lim_{n\to\infty} K_n(t) = t \qquad 0 \le t \le 1,$$

b) if $p = \frac{1}{2}$

$$\lim_{n\to\infty} K_n(t) = \begin{cases} 0 & 0 \le t \le \frac{1}{2} \\ 2t-1 & \frac{1}{2} \le t \le 1 \end{cases}$$

Proof: From (29) we have (for n odd)

$$K_n(\tfrac{1}{2}) = \sum_{k=1}^{(n+1)/2} \tfrac{1}{2}\alpha^{(k)} - \sum_{k=1}^{(n+1)/2} \alpha^{(k)} t^{(k)}. \tag{56}$$

But for n odd

$$\sum_{k=1}^{(n+1)/2} \alpha^{(k)} = 1 \tag{57}$$

and

$$\sum_{k=1}^{(n+1)/2} \alpha^{(k)} t^{(k)} = \sum_{k=1}^{(n+1)/2} \binom{n}{k-1}(1-p)^{k-1}(p)^{n-k+1}. \tag{57}$$

The sum on the right of (57) is just the probability that a random variable n, having a binomial distribution of mean $(n)(1-p)$ and variance $(n)(1-p)(p)$ will be less than $n/2$. If $p < \frac{1}{2}$, $(n)(1-p) > n/2$ and we may write (57) as

$$\begin{aligned}\sum_{k=1}^{(n+1)/2} \alpha^{(k)} t^{(k)} &= P_r\left\{x < \frac{n}{2}\right\} \\ &\leq P_r\{|\, x - (n)(1-p)\,| > (n)(\tfrac{1}{2} - p)\} \\ &\leq \frac{(p)(1-p)}{n(\frac{1}{2}-p)^2}.\end{aligned} \tag{58}$$

Where the last step is obtained from the Bienayme—Tchebycheff inequality. Now we use (56) and (57) in (55), let n approach infinity, and we obtain (for n odd and $p < \frac{1}{2}$)

$$\lim_{n\to\infty} K_n(\tfrac{1}{2}) = \tfrac{1}{2}. \tag{59}$$

Furthermore, since for a fixed r, $C_{n+1} \supset C_n$, (59) must also hold for even values of n. Finally, since $K_n(0) = 0$ and $K_n(1) = 1$ for all n, and $d/dt\, K_n(t)$ is a nondecreasing function of t, (59) proves part a) of the theorem.

Part b) is then proved simply by noting that if $p = \frac{1}{2}$, we have $r = 1$, and the only breakpoint of $K_n(t)$ is at $t = \frac{1}{2}$.

Lemma 5:

Let C_n be the nth semi-extension of the BSC C_1 and let C_1' be some other BSC. Then

a) $C_1' \supset C_n$ if and only if

$$t_{1,r'}^{(1)} \leq t_{n,r}^{(1)},$$

b) $C_n \supset C_1'$ if and only if

$$K_{n,r}(\tfrac{1}{2}) \geq K_{1,r'}(\tfrac{1}{2}).$$

Proof: From Lemma 1 we see that we need only prove

a) $$K_1'(t) \geq K_n(t) \quad \text{in} \quad [0, \tfrac{1}{2}]$$

if and only if

$$t_{1,r'}^{(1)} \leq t_{n,r}^{(1)},$$

b) $$K_n(t) \geq K_1'(t) \quad \text{in} \quad [0, \tfrac{1}{2}]$$

if and only if

$$K_{n,r}(\tfrac{1}{2}) \geq K_{1,r'}(\tfrac{1}{2}).$$

Note that $t_{1,r'}^{(1)}$ is the only breakpoint of $K_1^{(1)}(t)$ in $[0, \frac{1}{2}]$ so that by Lemma 3, the slope of $K_1^{(1)}(t)$ is unity in the interval $t_{1,r'}^{(1)} < t \leq \frac{1}{2}$. The slope of $K_n(t)$, however, is never greater than unity in this interval (again by Lemma 3). Parts a) and b) of Lemma 5 follow directly from these two facts.

On the Probability Density of the Output of a Low-Pass System When the Input is a Markov Step Process*

W. M. WONHAM†

Summary—**Forward equations are derived for the $(N+1)$-dimensional Markov process generated when a Markov step signal $\{s(t)\}$ is the input to an N^{th}-order system of the form $dX/dt = F(X; s)$. As examples, the joint probability densities of input and output are found for a symmetric three-level signal smoothed by an RC low-pass filter, and partial results are obtained for a doubly integrated Rice telegraph signal.**

* Received by the PGIT, January 8, 1960; revised manuscript received, June 2, 1960.

† Dept. of Engrg., University of Cambridge, Cambridge, Eng.

Introduction

If the input to a (possibly nonlinear) system is a Markov process, and the system belongs to a certain class of functionals, an equation for the characteristic function of the output distribution can be written using the methods of Darling and Siegert.[1]

[1] D. A. Darling and A. J. F. Siegert, "A systematic approach to a class of problems in the theory of noise and other random phenomena—I," IRE Trans. on Information Theory, vol. IT-3, pp. 32–37; March, 1957.

McFadden adapted the Darling-Siegert method to linear systems when the input is a Rice random telegraph signal,[2] and has developed techniques for other binary processes.[3] An alternative formulation of the problem, for Markov step inputs with a finite number of states, is given below in terms of forward equations for the joint densities of input and vector output. The step process, of which the Rice signal is a special case, arises in engineering applications as a model for discontinuous velocity signals.[4] Joint densities of input and output are of interest, for example, in the analysis of feedback control systems.

Markov Step Processes[5–7]

Let $\{s(t)\}$, $0 \leq t < \infty$ be a Markov process with a finite number of states $i = 1, \cdots, m$. To each state corresponds a bounded amplitude level s_i, where the s_i need not all be distinct. Stationary transition probabilities $p_{ij}(\tau)$,

$$p_{ij}(\tau) = Pr[s(t+\tau) = s_j \mid s(t) = s_i], \qquad \tau \geq 0, \tag{1}$$

are given by

$$\begin{aligned} P(\tau) &= [p_{ij}(\tau)] \\ &= \exp(\tau M) \\ &= I + \tau M + o(\tau) \quad \text{as} \quad \tau \to 0. \end{aligned} \tag{2}$$

Here I is the unit matrix, and $M = [\mu_{ij}]$ is a constant matrix such that

$$\mu_{ij} \geq 0, \qquad i \neq j, \tag{3}$$

$$\begin{aligned} \mu_{ii} &= -\sum_{j \neq i} \mu_{ij} \\ &\equiv -\mu_i. \end{aligned} \tag{4}$$

If $s(t) = s_i$, either a jump transition to s_j, $j \neq i$, occurs in a small time interval $(t, t+\tau)$, with probability $\mu_{ij}\tau + o(\tau)$; or no jump occurs, with probability $1 - \mu_i\tau + o(\tau)$. It can be shown[7] that the probability of no jump in an arbitrary interval $(t, t+\tau)$, conditional on $s(t) = s_i$, is $e^{-\mu_i\tau}$.

When limiting probabilities $p_{ij}(\infty) = p_j$ exist which are independent of the initial distribution, the p_j are given by

$$p_j\mu_j = \sum_{\substack{i=1 \\ i\neq j}}^{m} p_i\mu_{ij}$$

$$\sum_{j=1}^{m} p_j = 1.$$

(

A familiar example is the Rice random telegrap signal:[8] $s(t) = s_1$ or s_2, and

$$M = \mu\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix},$$

where μ is the expected number of jumps [zeros of $s(t$ in unit time.

In the following section, general equations are derive which apply to a fairly wide class of time-invariar systems with Markov step inputs. To illustrate th derivation, Example 1 is treated in complete detai

Forward Equations for the Joint Input-Output Process

Let the $s(t)$ process be the input to a time-invariar system with vector equation

$$\frac{dX}{dt} = U\Big[X; s(t)\Big],$$

(6

where

$$X(t) = [x^1(t), \cdots, x^N(t)], \qquad U = [u', \cdots, u^N].$$

(7

The system is "low-pass" in the sense that no derivativ of $s(t)$ appear in (6). It is assumed: 1) that if $s(t+\tau) = s$ $|\tau| \leq T$, then (6) yields a unique continuous pha trajectory

$$\begin{aligned} X(t+\tau) &= F[\tau; X(t), s_j] \\ &\equiv F_j[\tau; X(t)]; \ |\tau| \leq T, \quad j = 1, \cdots, m; \end{aligned}$$

(

and 2) that $X(t)$ is continuous at jump points of $s(t)$, i.e

$$|F_j(-\tau; X) - F_k(\tau; X)| \to 0 \quad \text{as} \quad \tau \to 0,$$

$$j, k = 1, \cdots, m$$

From (8), $X(t+\tau)$ is completely determined for $\tau \geq$ by $X(t)$, together with $s(t+\tau')$ for all τ', $0 \leq \tau' \leq$ Since $\{s(t)\}$ is a Markov process, the conditional distr bution of $s(t+\tau')$, $\tau' \geq 0$, given $s(t)$ and $s(t')$ for som $t' \leq t$, depends only on $s(t)$. Hence the conditional distr bution of the pair $[X(t+\tau), s(t+\tau)]$, given $[X(t), s($ and $[X(t'), s(t')]$ for some $t' \leq t$, depends only on t value of $[X(t), s(t)]$. That is, the joint process $\{X(t), s(t$ is Markov, continuous in the X component and m-valu in the s component.

Let $W_j(X, t)\, dx^1 \cdots dx^N = W_j(X, t)\, dX$ be the prob bility that $s(t) = s_j$ and $X(t) \in dX$ at X, conditior

[2] J. A. McFadden, "The probability density of the output of a filter when the input is a random telegraphic signal: differential equation method," IRE Trans. on Circuit Theory, vol. CT-6, pp. 228–233; May, 1959.

[3] J. A. McFadden, "The probability density of the output of an RC filter when the input is a binary random process," IRE Trans. on Information Theory, vol. IT-5, pp. 174–178; December, 1959.

[4] H. M. James, N. B. Nichols, and R. S. Phillips, "Theory of Servomechanisms," McGraw-Hill Book Co., Inc., New York, N. Y., ch. 6; 1947.

[5] A. Kolmogorov, "Analytische Methoden in der Wahrscheinlichkeitsrechnung," *Math. Ann.*, vol. 104, pt. 2, pp. 415–458; 1931. *Syst. Tech. J.*, vol. 23, article 2.7, p. 282; 1944.

[6] M. S. Bartlett, "An Introduction to Stochastic Processes," Cambridge University Press, Cambridge, Eng., ch. 3; 1955.

[7] J. L. Doob, "Stochastic Processes," John Wiley and Sons, Inc., New York, N. Y., ch. 6; 1953.

[8] S. O. Rice, "A mathematical analysis of random noise," I *Sys. Tech. J.*, vol. 23, article 2.7, p. 282; 1944.

on specified initial values $X(0)$, $s(0)$. Equations for the W_i may be derived as follows. Consider an arbitrary interval $(t, t+\tau)$, where $t, \tau > 0$. Then $X(t+\tau) = X$ and $s(t+\tau) = s_j$ if and only if: 1) $s(t) = s_j$, $X(t) = F_j(-\tau; X)$, and no jump of input occurs in $(t, t+\tau)$; or 2) for some τ' in $(0, \tau)$ and some $i \neq j$ the following conditions are satisfied: $s(t+\tau') = s_i$, $X(t+\tau') = F_j[-(\tau - \tau'); X]$, a jump $i \to j$ occurs in $d\tau'$ at $t + \tau'$, and no subsequent jump occurs in $(t+\tau', t+\tau)$. Adding the probabilities of events 1) and 2) yields

$$W_j(X, t+\tau) = e^{-\mu_j \tau} W_j[F_j(-\tau; X), t] \left| \frac{\partial F_j(-\tau; X)}{\partial X} \right|$$
$$+ \sum_{\substack{i=1 \\ i \neq j}}^{m} \int_0^{\tau} W_i[F_j(-(\tau - \tau'); X), t + \tau']$$
$$\cdot \left| \frac{\partial F_j(-(\tau - \tau'); X)}{\partial X} \right| \mu_{ij} e^{-\mu_j(\tau - \tau')}\, d\tau',$$
$$j = 1, \cdots, m. \quad (9)$$

The Jacobians allow for a change of volume element along the j trajectory (8) through X. The forward equations can be derived formally by expanding both sides of (9) to terms of $o(\tau)$, dividing by τ, and taking the limit $\tau \to 0$. From (6) and (8),

$$\begin{aligned} F_j(-\tau; X) &= X - \tau U(X; s_j) + o(\tau) \\ &\equiv X - \tau U_j(X) + o(\tau), \end{aligned} \quad (10)$$

where $U_j(X) = [u_j^1(X), \cdots, u_j^N(X)]$ is the velocity in phase space when $s = s_j$. From (10),

$$W_i[F_j(-\tau; X), t] = W_i(X, t)$$
$$- \tau \sum_{r=1}^{N} u_j^r(X) \frac{\partial W_i(X, t)}{\partial x^r} + o(\tau),$$
$$i, j = 1, \cdots, m. \quad (11)$$

Also,

$$\begin{aligned} \frac{\partial F_j(-\tau; X)}{\partial X} &= \det \frac{\partial}{\partial x^r} [x^s - \tau u_j^s(X) + o(\tau)] \\ &= \det [\delta_{rs} - \tau \frac{\partial u_j^s(X)}{\partial x^r} + o(\tau)] \\ &= 1 - \tau \sum_{r=1}^{N} \frac{\partial u_j^r(X)}{\partial x^r} + o(\tau). \end{aligned} \quad (12)$$

The integral in (9) expands simply to $\mu_{ij} W_i(X, t)\tau + o(\tau)$. On substituting (11), (12), and the last expression in (9), the limiting process yields the forward equations

$$\frac{\partial W_j(X, t)}{\partial t} + \operatorname{div} [U_j(X) W_j(X, t)]$$
$$= \sum_{\substack{i=1 \\ i \neq j}}^{m} \mu_{ij} W_i(X, t) - \mu_j W_j(x, t),$$
$$j = 1, \cdots, m; \quad (13)$$

where

$$\operatorname{div} [U_j W_j] \equiv \sum_{r=1}^{N} \partial(u_j^r W_j)/\partial x^r. \quad (14)$$

Summing (13) on j and using (4) gives the equation of continuity

$$\sum_{j=1}^{m} \left\{ \frac{\partial W_j(X, t)}{\partial t} + \operatorname{div} [U_j(X) W_j(X, t)] \right\} = 0. \quad (15)$$

In (13) and (15), $W_j(X, t)$ may be interpreted as the density of j-type gas at the point X, where $U_j(X)$ is the corresponding drift velocity, and $\mu_{ij} W_i$ is the rate of 'conversion' of i-type gas to j-type. Eq. (13) is analogous to Boltzmann's equation for a mixture of m gases[9] (with particle "conversion" in place of collisions), and might have been derived alternatively by an argument based on conservation of flow.

If the system (6) is one-dimensional ($N = 1$), (13), together with $s(0) = s_k$, $W_j(X, 0) = \delta_{jk}\delta(X - X_0)$, defines a hyperbolic (wave-type) initial value problem which has been solved for the Rice telegraph signal in special cases.[10,11] If $N \geq 2$, the system is no longer in general hyperbolic, but may be parabolic "in the broad sense."[12] Boundary values can be obtained by an iterative procedure as illustrated below in Example 2.

The equilibrium densities, when they exist, satisfy (13) with the time derivatives omitted. If $N = 1$, $X = x$, there results the ordinary system

$$\frac{d}{dx} [U_j(x) W_j(x)] = \sum_{\substack{i=1 \\ i \neq j}}^{m} \mu_{ij} W_i(x) - \mu_j W_j(x),$$
$$j = 1, \cdots, m. \quad (16)$$

In this case, integration of (15) yields

$$\sum_{j=1}^{m} U_j(x) W_j(x) = \text{constant} = 0. \quad (17)$$

That the constant of integration is zero is plausible on physical grounds, since the sum represents total "probability flow" along the x axis. For equilibrium densities to exist, the flow must vanish everywhere. On substituting (17) in (16) one of the unknown functions may be eliminated. Eq. (17) was suggested by the special case when $m = 2$ and the system is an RC low-pass filter. The latter result was pointed out to the writer by Dr. McFadden,[13] who derived it using a level-crossing argument.[14]

[9] J. H. Jeans, "Kinetic Theory of Gases," Cambridge University Press, Cambridge, Eng., ch. 9; 1940.
[10] S. Goldstein, "On diffusion by discontinuous movements and on the telegraph equation," *Quart. J. Mech. and Appl. Math.*, vol. 4, pp. 129–156; 1951.
[11] W. M. Wonham, "Transition probability densities of the smoothed random telegraph signal," *J. Electronics and Control*, vol. 6, pp. 376–384; 1959.
[12] I. G. Petrovsky, "Lectures on Partial Differential Equations," Interscience Publishers, Inc., New York, N. Y., ch. 1; 1954.
[13] J. A. McFadden, private communication.
[14] J. A. McFadden, "The probability density of the output of a filter when the input is a random telegraphic signal." 1959 IRE National Convention Record, pt. 4, pp. 164–169. [*Cf.* (22) and (23).]

Example 1: First-Order RC Filtering of a Three-Level Step Signal

Let $\{s(t)\}$ be a symmetric three-level step signal, with $s_1 = -s_3 = 1$, and $s_2 = 0$. The transition matrix is

$$P(\tau) = [p_{ij}(\tau)] = I - \tau M + o(\tau), \tag{18}$$

where

$$M = [\mu_{ij}] = \mu \begin{bmatrix} -1 & 1-\alpha & \alpha \\ \beta & -2\beta & \beta \\ \alpha & 1-\alpha & -1 \end{bmatrix};$$

$$\mu, \beta > 0; \qquad 0 \leq \alpha < 1. \tag{19}$$

From (5) and (19), the limiting probabilities are

$$p_1 = p_3 = \beta/(1 - \alpha + 2\beta);$$
$$p_2 = (1 - \alpha)/(1 - \alpha + 2\beta). \tag{20}$$

On calculation of $P(t) = e^{Mt}$, the covariance function is found to be

$$R(\tau) \equiv E\{s(t)s(t + \tau)\} = \sum_{i,j} s_i p_i p_{ij}(\tau) s_j = 2p_1 e^{-\mu(1+\alpha)|\tau|}. \tag{21}$$

The equilibrium densities will be found when $s(t)$ is the input to a first-order low-pass RC filter with RC = 1 and differential equation

$$\frac{dx}{dt} = s(t) - x. \tag{22}$$

If $s(t') = \text{constant} = s_j$ in the interval $t \leq t' \leq t + \tau$, (22) may be solved for $x(t)$ in terms of $x(t + \tau)$:

$$x(t) = e^{\tau}[x(t + \tau) - s_j] + s_j. \tag{23}$$

Let $W_j(x, t)\, dx$ be the joint probability that $s(t) = s_j$ and $x \leq x(t) \leq x + dx$ [conditional on specified initial values $s(0)$, $x(0)$]. For an arbitrary interval $(t, t + \tau)$, one has $s(t + \tau) = s_j$ and $x(t + \tau) = x$, if and only if: 1) $s(t) = s_j$, $x(t) = x_1 \equiv e^{\tau}(x - s_j) + s_j$, and no jump of input occurs in $(t, t + \tau)$; or 2) for some $i \neq j$ and some τ' in $(0, \tau)$ the following conditions are satisfied: $s(t + \tau') = s_i$, $x(t + \tau') = x' \equiv e^{(\tau-\tau')}(x - s_j) + s_j$, a jump $i \rightarrow j$ occurs in $d\tau'$ at $t + \tau'$, and no subsequent jump occurs in $(t + \tau', t + \tau)$. Adding the probabilities of events 1) and 2),

$$W_j(x, t + \tau) = e^{-\mu_j\tau} W_j(x_1, t)(dx_1/dx) + \sum_{i \neq j} \int_0^{\tau} W_i(x', t + \tau')(dx'/dx) e^{-\mu_j(\tau-\tau')} \mu_{ij}\, d\tau'. \tag{24}$$

On differentiating both sides of (24) with respect to τ, and setting $\tau = 0$, there results the following special case of (13):

$$\frac{\partial W_j(x, t)}{\partial t} + \frac{\partial}{\partial x}[(s_j - x)W_j(x, t)] = \sum_{i \neq j} \mu_{ij} W_i(x, t) - \mu_j W_j(x, t), \quad j = 1, 2, 3. \tag{25}$$

Eqs. (25) are the forward equations for the joint input output process $\{s(t), x(t)\}$.

It is reasonable to assume that equilibrium densitie $W_j(x)$ exist in the limit $t \rightarrow \infty$. Since $|s(t)| \leq 1$, it i seen from (23) that $|x(t)| \leq 1$, so $W_j(x) = 0$, $|x| > 1$ Dropping the time variable in (25) and substituting values of the s_j and μ_{ij} gives

$$(d/dx)(1 - x)W_1(x) = \mu(-W_1 + \beta W_2 + \alpha W_3), \tag{26a}$$

$$-(d/dx)xW_2(x) = \mu[(1 - \alpha)W_1 - 2\beta W_2 + (1 - \alpha)W_3], \tag{26b}$$

$$-(d/dx)(1 + x)W_3(x) = \mu[\alpha W_1 + \beta W_2 - W_3], \quad |x| \leq 1. \tag{26c}$$

By definition of $W_j(x)$, one has $\int_{-1}^{1} W_j(x)\, dx = p_j$ Integrating (26a) from -1 to 1 and using (20) yield

$$[(1 - x)W_1(x)]_{-1}^{1} = 0.$$

As $x \rightarrow +1$, $(1 - x)W_1(x) \rightarrow 0$, otherwise W_1 would not be integrable. Hence

$$W_1(-1) = 0; \tag{27a}$$

and similarly

$$W_2(\pm 1) = W_3(+1) = 0. \tag{27b}$$

On integrating the sum of (26), there now follows

$$(1 - x)W_1(x) - xW_2(x) - (1 + x)W_3(x) = \text{constant} = 0. \tag{28}$$

That the constant of integration is zero is seen by lettin $x \rightarrow \pm 1$.

Eqs. (26) with end conditions (27) can be solved i terms of the hypergeometric series

$$H_1(z) = {}_2F_1\left[\frac{1 + \mu(1 + \alpha - 2\beta)}{2}, \frac{\mu(1 - \alpha)}{2}; \mu; z\right], \tag{29}$$

$$H_2(z) = {}_2F_1\left[\frac{1 + \mu(1 + \alpha - 2\beta)}{2}, 1 + \frac{\mu(1 - \alpha)}{2}; 1 + \mu; z\right]. \tag{30}$$

If $|x| > 1$, $W_j(x) = 0$. For $|x| \leq 1$, the results ar

$$W_1(x) = C(1 + x)(1 - x^2)^{\mu-1} \cdot \left[H_1(1 - x^2) - \left(\frac{1 - \alpha}{2}\right)(1 - x)H_2(1 - x^2)\right], \tag{31a}$$

$$W_2(x) = C(1 - \alpha)(1 - x^2)^{\mu} H_2(1 - x^2), \tag{31b}$$

$$W_3(x) = W_1(-x), \tag{31c}$$

$$\begin{aligned} W(x) &\equiv W_1 + W_2 + W_3, \\ &= 2C(1 - x^2)^{\mu-1}H_1(1 - x^2). \end{aligned} \tag{32}$$

The normalizing constant C is

$$C = \frac{\Gamma[(\mu/2)(1 - \alpha + 2\beta)]\Gamma[(1 + \mu(1 + \alpha))/2]}{2\sqrt{\pi}\,\Gamma(\mu)\Gamma(\mu\beta)}. \tag{33}$$

The densities simplify for special values of the parameters. For example, if $1 + \mu(1 + \alpha - 2\beta) = -2K$, $K = 0, 1, 2, \cdots$, the series (29), (30) stop after $K + 1$ terms. In this case, the total output density $W(x)$ can be written in terms of a Jacobi polynomial[15] of Kth degree:

$$\begin{aligned} W(x) = &\frac{K!\,\Gamma(2\mu + 2K)\Gamma(\mu\beta - K)}{2^{2(\mu+K)-1}\Gamma(\mu\beta)[\Gamma(\mu + K)]^2} \\ &\cdot(1 - x^2)^{\mu-1}P_K^{(\mu-1,1/2-\mu\beta)}(2x^2 - 1), \quad |x| \leq 1. \end{aligned} \tag{34}$$

It is interesting to note that if $K = 0$, *i.e.*, $2\beta = \mu^{-1} + 1 + \alpha$, $W(x)$ is identical with the output density obtained when the input is a Rice telegraph signal.[16] The three-state signal actually reduces to the latter only when $\alpha \to 1$.

Another class of simpler distributions is obtained when μ is small, and also $2\beta = \mu^{-1} - (1 - \alpha)$. Then

$$\begin{aligned} W(x) = &\frac{\Gamma[(1 + \mu + \mu\alpha)/2]}{\Gamma(\mu)\Gamma[(1 - \mu + \mu\alpha)/2]} \\ &\cdot|x|^{-\mu(1-\alpha)}(1 - x^2)^{\mu-1}. \end{aligned} \tag{35}$$

The moments of $W(x)$ are readily obtained from (32), since integration leads to a ${}_2F_1$ series with argument unity, which can be summed. It can then be verified that for arbitrarily-fixed α, β, and large μ, the distribution of $x\sqrt{(1 + \mu + \mu\alpha)/2p_1}$ is asymptotically Gaussian with zero mean and unit variance.

Example 2: Doubly Integrated Rice Telegraph Signal

Let $\{s(t)\}$ be the Rice telegraph signal with $s_1 = +1$, $s_2 = -1$, and x, y the output of the first and second integrator, respectively. Eqs. (6) and (8) for this case are

$$(d/dt)(x, y) = (s, x); \tag{36}$$

$$\begin{aligned} (x, y) &= F_j(t; x_0, y_0) \\ &= (x_0 + s_j t, x_0 t + y_0 + s_j t^2/2). \end{aligned} \tag{37}$$

Eqs. (13) become the parabolic system

$$\begin{aligned} \partial W_1/\partial t + \partial W_1/\partial x + x\partial W_1/\partial y + \mu W_1 &= \mu W_2 \\ \partial W_2/\partial t - \partial W_2/\partial x + x\partial W_2/\partial y + \mu W_2 &= \mu W_1. \end{aligned} \tag{38}$$

The initial values will be taken as $s(0) = s_1 = +1$, $x(0) = y(0) = 0$. If $s(0) = -1$, it is clear from symmetry that W_1, W_2 are to be interchanged, and (x, y) replaced by $(-x, -y)$. For arbitrary initial values (x_0, y_0), (x, y) is replaced by $(x - x_0, y - x_0 t - y_0)$.

Let $w^{(q)}(x, y, t)$, $q = 0, 1, 2, \cdots$, be the probability density of (x, y) at time t together with the probability that exactly q zeros of input occur in $(0, t)$, conditional on the initial values. Since $q = 0$ with probability $e^{-\mu t}$,

$$w^{(0)}(x, y, t) = e^{-\mu t}\,\delta(x - t)\,\delta(y - t^2/2). \tag{39}$$

If one input zero occurs in $(0, t)$, and $s(0) = -1$, the locus of terminal points at time t is found from (37) to be

$$y = \psi_1(x, t) = [(t + x)^2 - 2t^2]/4, \qquad |x| \leq t. \tag{40}$$

Similarly, if $s(0) = +1$, the new locus is

$$y = \psi_2(x, t) = [-(t - x)^2 + 2t^2]/4, \qquad |x| \leq t. \tag{41}$$

Using (37), (40), and (41), it can be shown by induction on the number of input zeros in $(0, t)$ that the domain of the W_j is the conoid

$$C: \psi_1(x, t) \leq y \leq \psi_2(x, t), \qquad |x| \leq t, \qquad t \geq 0. \tag{42}$$

The characteristic surfaces $y = \psi_j$ are shown in Fig. 1.

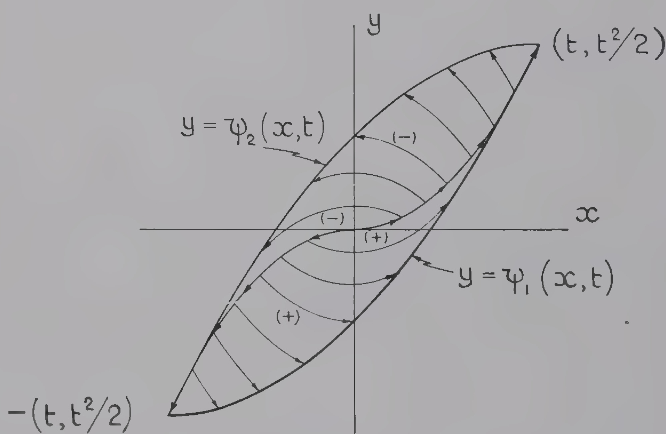


Fig. 1—Domain of probability densities $W_j(x, y, t)$ for a doubly-integrated Rice telegraph signal $s(t) = \pm 1$.
$dx/dt = s(t)$; $dy/dt = x(t)$
$|x| \leq t$; $\psi_1(x,t) \leq y \leq \psi_2(x, t)$.
$(+)$, $(-)$ denote trajectories corresponding to $s(t) = +1, -1$.

To find boundary values of the W_j on these surfaces it is sufficient to compute the first few functions $w^{(q)}(x, y, t)$. By an argument similar to the one which led to (9), one can write

$$w^{(q)}(x, y, t) = \mu \int_0^{\tau_j} e^{-\mu\tau} w^{(q-1)}[F_j(-\tau; x, y); t - \tau]\, d\tau, \qquad q = 1, 2, \cdots. \tag{43}$$

Since $s(0) = s_1$, $j = 1$ or 2 in the integrand, depending

[15] A. Erdélyi, *et al.*, "Higher Transcendental Functions," McGraw-Hill Book Co. Inc., New York, N. Y., article 10.8.; 1953.
[16] W. M. Wonham, and A. T. Fuller, "Probability densities of the smoothed random telegraph signal," *J. Electronics and Control*, vol. 4, pp. 567–576; 1958. [See (27).]

on whether q is even or odd. The integration is taken over all previous instants $t - \tau$ at which the $(q-1)$st zero may have occurred. The upper limit τ_j is therefore defined by the intersection of the trajectory $(x', y') = F_j(-\tau; x, y)$ with the boundary surface $y' = \psi_2(x', t - \tau)$ (if $j = 1$), or $y' = \psi_1(x', t - \tau)$ (if $j = 2$). From (37), (40) and (41) there results

$$\begin{aligned} \tau_1 &= [\psi_2(x, t) - y]/(t - x), \\ \tau_2 &= [y - \psi_1(x, t)]/(t + x). \end{aligned} \tag{44}$$

From (39), (43), and (44) one may now derive

$$w^{(1)}(x, y, t) = (\mu e^{-\mu t}/2)\, \delta(y - \psi_2), \tag{45a}$$

$$w^{(2)}(x, y, t) = \mu^2 e^{-\mu t}/2(t - x), \tag{45b}$$

$$w^{(3)}(x, y, t) = (\mu^3 e^{-\mu t}/4) \log\left[\frac{\psi_2 - \psi_1}{\psi_2 - y}\right], \tag{45c}$$

$$w^{(q)}(x, \psi_j, t) = 0; \quad j = 1, 2; \quad q = 4, 5, \cdots, \tag{46}$$

Since $s(t) = s_1$ or s_2, depending on whether q is even or odd, the space densities W_j are given by

$$\begin{aligned} W_1 &= \sum_{q=1}^{\infty} w^{(2q)}, \\ W_2 &= \sum_{q=1}^{\infty} w^{(2q+1)}. \end{aligned} \tag{47}$$

Thus the boundary values are, finally,

$$\begin{aligned} &W_1(x, \psi_j, t) = w^{(2)}(x, \psi_j, t), \\ &W_2(x, y, t) \sim w^{(3)}(x, y, t) \quad \text{as} \quad y \to \psi_j. \end{aligned} \tag{48}$$

Eqs. (38) and (48) have been solved up to a Laplace transform with respect to y. On setting

$$\bar{W}_j(x, p, t) = \int_{\psi_1}^{\psi_2} e^{-py} W_j(x, y, t)\, dy,$$

(38) and (48) become

$$\begin{aligned} &\partial\bar{W}_1/\partial t + \partial\bar{W}_1/\partial x + (\mu + px)\bar{W}_1 \\ &\quad - (\mu^2/2) \exp(-\mu t - p\psi_2) = \mu\bar{W}_2 \\ &\partial\bar{W}_2/\partial t - \partial\bar{W}_2/\partial x + (\mu + px)\bar{W}_2 = \mu\bar{W}_1, \end{aligned} \tag{49}$$

$$\bar{W}_1(x, p, t) = \begin{cases} (\mu^2 t/2) \exp(-\mu t - pt^2/2), & x = t \\ 0 & x = -t \end{cases}$$

$$\bar{W}_2(x, p, t) = 0, \qquad x = \pm t. \tag{50}$$

The transformed system is hyperbolic, with boundary values given on the characteristics $x = \pm t$, $t \geq 0$, and can be solved by standard methods. The results may be written in terms of confluent hypergeometric functions:

$$\begin{aligned} \bar{W}_1(x, p, t) &= (\mu^2/4)(t + x) \exp(-\mu t - p\psi_2) \\ &\quad \cdot \{{}_1F_1\, [1 + \mu^2/2p; 2; p(t^2 - x^2)/2]\} \end{aligned} \tag{51}$$

$$\begin{aligned} \bar{W}_2(x, p, t) &= (\mu/2) \exp(-\mu t - p\psi_2) \\ &\quad \cdot \{{}_1F_1[\mu^2/2p; 1; p(t^2 - x^2)/2] - 1\}. \end{aligned} \tag{5}$$

If $s(0) = -1$, the new transformed space densitie are obtained by interchanging $\bar{W}_1$, $\bar{W}_2$ and replacin (x, p) by $(-x, -p)$. The total space density condition on

$$\left.\begin{aligned} &x(0) = y(0) = 0 \\ &s(0) = +1 \quad \text{or} \quad -1 \quad \text{with equal probability} \end{aligned}\right\} \tag{5}$$

is, therefore,

$$\begin{aligned} \bar{W}(x, p, t) &\equiv \tfrac{1}{2}[\bar{W}_1(x, p, t) + \bar{W}_1(-x, -p, t) \\ &\quad + \bar{W}_2(x, p, t) + \bar{W}_2(-x, -p, t)] \\ &= (\mu/4)e^{-\mu t - p\psi_2}\{2\, {}_1F_1[\mu^2/2p; 1; p(t^2 - x^2)/2] \\ &\quad + [\mu t + p(t^2 - x^2)/2]\, {}_1F_1[1 + \mu^2/2p; 2; p(t^2 - x)/2] \\ &\quad - e^{p(t^2 - x^2)/2} - 1\}. \end{aligned} \tag{5}$$

In (54), Kummer's transformation[17] was used to evaluat $\bar{W}_j(-x, -p, t)$.

The marginal density for $x(t)$ alone can be found o integrating $w^{(0)}(\pm x, \pm y, t)$, $w^{(1)}(\pm x, \pm y, t)$ with respec to y and setting $p = 0$ in (54); this result has been dis cussed by McFadden.[14] The expectation of $y(t)$ conditiona on $x(t) = x$ and initial values (53) is simply

$$\begin{aligned} E\{y \mid x\} &= (\psi_1 + \psi_2)/2 \\ &= tx/2. \end{aligned} \tag{5}$$

The conditional variance of $y(t)$ can be obtained fro (45a), (55), and the coefficient of $p^2/2$ in (54). The exac result is somewhat involved; however, two approximatio are

$$E\{y^2 \mid x\} - [E\{y \mid x\}]^2 = \begin{cases} \dfrac{(t^2 - x^2)^2}{16}\left(1 - \dfrac{\mu t}{3}\right) + o(\mu t) \quad \text{as} \quad \mu t \to 0. \\ \dfrac{(t^2 - x^2)^{3/2}}{12\mu} + 0\left(\dfrac{1}{\mu^2}\right); \; t, x \quad \text{fixed}; \quad \mu \to \infty. \end{cases} \tag{5}$$

If $|x(t)| < t$ and $\mu t \to 0$, only the boundary densiti $w^{(1)}(\pm x, \pm y, t)$ contribute to the conditional varianc For $\mu t > 0$ and small the only first-order term aris from $w^{(2)}(\pm x, \pm y, t)$. When μ is large, the absolu variances of $x(t)$ and $y(t)$ are both small, and the co ditional y-distribution sharply peaked at the mean.

Acknowledgment

The writer wishes to thank Dr. A. T. Fuller of Ca bridge University for useful criticism, and Dr. J. A. M Fadden, Purdue University, for helpful corresponden and the early communication of footnotes 2, 3, and 1 Financial assistance was received from the National R search Council of Canada.

[17] Erdélyi, *op. cit.*, article 6.3, (7).

Generation of a Sampled Gaussian Time Series Having a Specified Correlation Function*

MORRIS J. LEVIN†, MEMBER, IRE

Summary—**A computationally convenient method is presented for simulating a sequence of sampled values of a stationary Gaussian process having a specified correlation function or power spectrum. Z-transform theory is applied to provide a simple recursive formula for generating the values from a set of independent Gaussian random variables.**

Introduction

THE PROBLEM considered here is that of generating numerically from a table of independent Gaussian deviates[1] a finite sequence of random variables having the statistical properties of a set of uniformly spaced samples of a stationary Gaussian process with a specified correlation function or power spectrum. Such sequences are useful for the digital simulation of noise problems and for other purposes. Stein and Storer[2] have discussed some applications of these sequences and also considered several methods for generating them. Two approximate methods were suggested but the effects of the approximations were not stated. In addition, an exact procedure was presented which, for the computation of N sample values, requires the computation of the eigenvectors and eigenvalues of the covariance matrix of the sequence, which is of order N. Marsaglia[3] pointed out that equivalent results can be obtained by the much simpler procedure of factoring the covariance matrix into two triangular matrices and then multiplying the sequence of independent samples by one of these matrices. However, for large N, say 100, this is still a formidable computational problem. This paper provides a greatly simplified procedure for accomplishing this task when the desired sequence is derived from a stationary process with known correlation function or power spectrum. In this case the necessity for factoring the covariance matrix is eliminated and after some relatively short preliminary calculations the sample sequence is given by a simple linear recursive formula which is applicable for N arbitrarily large.

It is assumed that the process which is sampled has mean zero and a power spectrum $P(\omega^2)$ which is rational and of order K in ω^2. The method described here requires either the correlation function or the power spectrum of the sampled time series. For simplicity we assume that this sampled time series is rescaled in time so that its sampling interval is unity. Then if the correlation function $p(\tau)$ of the continuous process is given, the corresponding rescaled sampled correlation function is

$$\phi(m) = p(mT) \qquad m = 0, \pm 1, \pm 2, \cdots \tag{1}$$

where T is the sampling interval. If the power spectrum $P(\omega^2)$ is given, then the corresponding sampled power spectrum can be determined by the method described by Ragazzini and Franklin.[4]

Notation

The following notation is employed:

$z = e^{i\omega}$

$\phi(m) =$ the correlation function of the sampled time series to be simulated.

$\Phi(z) = \sum_{m=-\infty}^{\infty} \phi(m) z^{-m} =$ the sampled power spectrum corresponding to $\phi(m)$, *i.e.*, the z transform of $\phi(m)$.

$u(n),\ n \geq 0 =$ a sequence of independent Gaussian deviates with mean zero and variance one.

$v(i),\ 0 \leq i \leq K-1 =$ an auxiliary set of independent Gaussian deviates with mean zero and variance one.

$y(n),\ n \geq 0 =$ the random sequence having correlation function $\phi(m)$ which is generated from $u(n)$.

$\xi_i,\ 0 \leq i \leq K-1 =$ an auxiliary sequence of random variables which is generated from $v(i)$.

$R_{ij} =$ covariance of ξ_i and ξ_j.

$H(z) = \dfrac{a_0 + a_1 z^{-1} + \cdots + a_K z^{-K}}{1 + b_1 z^{-1} + \cdots + b_K z^{-K}}, \quad (b_K \neq 0) =$ the pulse transfer function of a stable linear sampled-data filter the output of which is a time series having correlation function $\phi(m)$ when the input is a sequence of independent random variables with variance one.

$h(n) =$ the impulse response of the filter $H(z)$.

$\Gamma =$ the unit circle in the complex plane.

$E =$ statistical expectation (mean) operator.

Description of the Procedure

The basic idea of the procedure described here is to calculate the transfer function $H(z)$ of a linear filter which

* Received by the PGIT, February 15, 1960.

† RCA Missile Electronics and Controls Div., Burlington, Mass.

[1] M. E. Muller, "A comparison of methods for generating normal deviates on digital computers," *J. Assoc. for Computing Mach.*, vol. 6, pp. 376–383; July, 1959.

[2] S. Stein and J. E. Storer, "Generating a Gaussian sample," IRE Trans. on Information Theory, vol. IT-2, pp. 87–90; June, 1956.

[3] G. Marsaglia, "A note on the construction of a multivariate normal sample," IRE Trans. on Information Theory, vol. IT-3, p. 149; June, 1957.

[4] J. R. Ragazzini and G. F. Franklin, "Sampled-Data Control Systems," McGraw-Hill Book Co., Inc., New York, N. Y. p. 259; 1958.

would convert white noise into noise with the specified correlation function $\phi(m)$, and then to use $H(z)$ expressed as a recursion relationship to compute $y(n)$ from $u(n)$. For the filter $H(z)$ to generate a stationary random output, it must be operating in the "steady state;"[5,6] *i.e.*, it must have had an input of white noise for all $n \geq -\infty$. However, for computational purposes the input sequence must begin at $n = 0$. The procedure described here provides the correct "initial conditions" by generating K values of $y(n)$, $0 \leq n \leq K - 1$, by a special method so that these K values together with the corresponding values of $u(n)$ have the same covariance matrix as if the filter were operating in the steady state. This is done by replacing the effect of the input sequence for $n < 0$ by K auxiliary random variables, ξ_i, which are generated from the auxiliary independent random variables $v(i)$. The simplicity of the approach rests on the fact that in a practical problem K will seldom exceed two or three. Then the remaining values of $y(n)$, $n \geq K$, can be computed recursively. The steps in this procedure are described below.

Step 1: *Determination of* $H(z)$

If $\Phi(z)$ is not specified, it must be calculated by taking the z transform of $\phi(m)$. Then $H(z)$ is determined by[7]

$$\Phi(z) = H(z)H(z^{-1}). \tag{2}$$

Since $H(z)$ is assumed to be stable it has all its poles within Γ and $H(z^{-1})$ has all its poles outside. Therefore $H(z)$ is found by factoring $\Phi(z)$ and associating appropriate poles and zeros. Since the zeros of $H(z)$ are not restricted to be within Γ, there are alternative forms for $H(z)$, depending upon which zeros of $\Phi(z)$ are associated with $H(z)$ and which with $H(z^{-1})$.

Step 2: *Determination of* $h(n)$

In general,

$$H(z) = \sum_{n=0}^{\infty} h(n)z^{-n}. \tag{3}$$

$h(n)$ is easily obtained by expanding $H(z)$ by long division. Only the first K values of $h(n)$ are required here.

Step 3: *Determination of* $y(n)$, $0 \leq n \leq K - 1$

A. Formulation: The output of a linear sampled-data filter is a linear function of the present and all previous input values,

$$y(n) = \sum_{m=0}^{\infty} h(m)x(n - m) \quad (4$$

where $x(n)$ represents the input to the filter. Then we ca write, for example,

$$y(0) = \sum_{m=0}^{\infty} h(m)x(-m) = h(0)x(0) + \xi_0 \quad (5$$

where

$$\xi_0 = \sum_{m=1}^{\infty} h(m)x(-m). \quad (6$$

In general, we can take

$$y(n) = \sum_{m=0}^{n} h(m)x(n - m) + \xi_n, \qquad n \leq K - 1 \quad (7$$

where

$$\xi_n = \sum_{m=1}^{\infty} h(m + n)x(-m) \quad (8$$

represents the influence of all $x(n)$ for $n < 0$. The virtu of this representation is that if the correlation functio of $x(n)$ is known, then the covariance matrix of the can be determined. With this matrix available, values o the ξ_i can be generated having the appropriate statistic properties so as to simulate the effect of all $x(n)$ for $n <$

B. Covariance Matrix of the ξ_i: For the situation co sidered here, the $x(n)$ are statistically independent so tha for example,

$$\begin{aligned} R_{00} = \operatorname{Var} \xi_0 &= E\left[\sum_{m=1}^{\infty} h(m)x(-m)\right]^2 \\ &= \sum_{m=0}^{\infty} h^2(m) - h^2(0) \\ &= \phi(0) - h^2(0). \end{aligned} \quad (9$$

From z-transform theory,[7] $\phi(0)$ can be determined fro $\Phi(z)$ by taking

$$\phi(0) = \sum \text{residues of poles within } \Gamma \text{ of } \frac{\Phi(z)}{z}. \quad (1$$

In general, we find

$$\begin{aligned} R_{ij} &= \operatorname{cov}(\xi_i, \xi_j) \\ &= E \sum_{m=1}^{\infty} \sum_{p=1}^{\infty} h(m + i)h(p + j)x(-m)x(-p) \\ &= \sum_{m=1}^{\infty} h(m + i)h(m + j) \\ &= \sum_{k=i+1}^{\infty} h(k)h(k + j - i) \\ &= \phi(j - i) - \sum_{m=0}^{i} h(m)h(m + j - i) \end{aligned} \quad (1$$

[5] B. Friedland, "Least squares filtering and prediction of nonstationary sampled data," *Information and Control*, vol. 4, pp. 297–313; December, 1958.

[6] D. G. Lampard, "The response of linear networks to suddenly applied random noise," IRE Trans. on Circuit Theory, vol. CT-2, pp. 49–57; March, 1955.

[7] R. H. Barker, "The pulse transfer function and its application to sampling servo systems," *Proc. IEE*, vol. 99, pt. 4, Monograph 43; July 15, 1952.

and

$$\phi(m) = \sum \text{residues within } \Gamma \text{ of } [z^{|m|-1}\Phi(z)]. \qquad (12)$$

C. Determination of the ξ_i: After having calculated the covariance matrix of the K random variables ξ_i, appropriate sample values for these variables can be obtained by linearly transforming K auxiliary independent Gaussian deviates, $v(i)$. A convenient method for doing this takes the form

$$\xi_i = \sum_{j=0}^{i} c_{ij} v(j). \qquad (13)$$

The values of the c_{ij} can be found by the method described by Marsaglia.[3] For convenience we list here recursive formulas for the first few c_{ij}, which are sufficient if $K \leq 3$.

$$c_{00} = \sqrt{R_{00}},$$

$$c_{10} = \frac{R_{10}}{c_{00}}, \quad c_{11} = \sqrt{R_{11} - c_{10}^2} \qquad (14)$$

$$c_{20} = \frac{R_{20}}{c_{00}}, \quad c_{21} = \frac{R_{21} - c_{10}c_{20}}{c_{11}}, \quad c_{22} = \sqrt{R_{22} - c_{20}^2 - c_{21}^2}.$$

Note that only these K values of the ξ_i are required no matter how large N is.

Step 4: Determination of $y(n)$ for $n \geq K$

The transfer function

$$H(z) = \frac{Y(z)}{U(z)} = \frac{a_0 + a_1 z^{-1} + \cdots + a_K z^{-K}}{1 + b_1 z^{-1} + \cdots + b_K z^{-K}} \qquad (15)$$

is equivalent to the recursive relation

$$y(n) = -b_1 y(n-1) - \cdots - b_K y(n-K) + a_0 u(n) + a_1 u(n-1) + \cdots + a_K u(n-K). \qquad (16)$$

The procedure described in the previous sections provides K consecutive pairs of corresponding values of $y(n)$ and $u(n)$ having the same statistical properties as if they were produced by steady-state operation of a filter having the transfer function $H(z)$. Therefore, any desired additional number of values of $y(n)$ can be generated by applying the recursive relation (16) starting with these initial values. We have, in a sense, provided the correct initial conditions, so that the filter now behaves as if the sequence $u(n)$ had been applied to it for all $n \geq -\infty$ and not just $n \geq 0$. It should be noted that the $y(n)$ sequence which is generated is Gaussian since every value is obtained as a linear combination of Gaussian variables.

Example 1:

Given

$$p(\tau) = e^{-\alpha|\tau|};$$

then from (1),

$$\phi(m) = e^{-\alpha T|m|}.$$

The z transform of $\phi(m)$ is obtained by taking the sum of the individual z transforms of the parts for $m \geq 0$ and $m < 0$. Letting $A = e^{-\alpha T}$ we have

$$\Phi(z) = \frac{1}{1 - Az^{-1}} + \frac{1}{1 - Az} - 1 = \frac{\sqrt{1 - A^2}}{1 - Az^{-1}} \cdot \frac{\sqrt{1 - A^2}}{1 - Az}.$$

From (2) and (3)

$$H(z) = \frac{\sqrt{1 - A^2}}{1 - Az^{-1}} = \sqrt{1 - A^2}\,[1 + Az^{-1} + A^2 z^{-2} + \cdots].$$

Then

$$h(0) = \sqrt{1 - A^2},$$

so from (9) and (14)

$$c_{00} = \sqrt{\phi(0) - h^2(0)} = A,$$

and from (13)

$$\xi_0 = c_{00}\, v(0) = Av(0).$$

Then by (5)

$$y(0) = \sqrt{1 - A^2}\, u(0) + Av(0).$$

Since $u(0)$ and $v(0)$ are independent and their values do not enter the expression for $y(n)$ for $n \geq 1$, $y(0)$ can be generated more simply from a single random variable having the appropriate variance, by taking

$$y(0) = [(\sqrt{1 - A^2})^2 + A^2]^{1/2} u(0) = u(0).$$

Finally for $n \geq 1$, from (16)

$$y(n) = \sqrt{1 - A^2}\, u(n) + Ay(n-1).$$

Example 2:

Given

$$\Phi(z) = \frac{64z^2}{8z^4 + 54z^3 + 101z^2 + 54z + 8}.$$

By factoring we find

$$H(z) = \frac{z^2}{(z + \frac{1}{2})(z + \frac{1}{4})} = \frac{1}{1 + 0.75z^{-1} + 0.125z^{-2}}.$$

By long division $h(0) = 1$ and $h(1) = -0.75$. Referring

to (10) and (12)

$$\phi(0) = \sum \text{residues within } \Gamma \text{ of} \left[\frac{\Phi(z)}{z}\right]$$
$$= 64/35$$
$$\phi(1) = \sum \text{residues within } \Gamma \text{ of } \Phi(z)$$
$$= -128/105.$$

Therefore from (11)

$$R_{00} = \frac{29}{35}, \qquad R_{10} = \frac{-197}{420}, \qquad R_{11} = \frac{149}{560}.$$

Substituting in (14),

$$c_{00} = 0.910, \qquad c_{10} = -0.515, \qquad c_{11} = 0.023,$$

and from (7), (13), and (16),

$$y(0) = u(0) + 0.91\, v(0),$$
$$y(1) = u(1) - 0.75\, u(0) - 0.515\, v(0) + 0.023\, v(1),$$
$$y(n) = -0.75\, y(n-1) - 0.125\, y(n-2) + u(n) \quad n \geq 2$$

Conclusions

The procedure presented here provides an exact an computationally practical method for solving the problem considered. It might appear that the sequence of step described for obtaining $y(n)$ for $0 \leq n \leq K - 1$ coul be consolidated into explicit formulas, but no method o doing this, which results in any simplification, has bee found.

A Note on the Local Structure of Shot Noise*

RICHARD A. SILVERMAN†, SENIOR MEMBER, IRE

Summary—It is shown how to construct shot noise which, like the turbulent velocity field, is quite accurately univariate normal, but exhibits marked departures from bivariate normality at close ranges.

It is well known that the random velocity field produced by wind tunnel turbulence at high Reynolds numbers has the following statistical structure: Let the x axis be directed downstream beyond the grid generating the turbulence and let $u_1(x, y, z)$ be the downstream component of the random velocity field, measured at any point of the flow. Then, to a good approximation, $u_1(x, y, z)$ is a normal random variable, *i.e.*, the velocity field is quite accurately univariate normal. If the field were bivariate normal as well, then writing $\delta u(r) = u_1(x + r, y, z) - u_1(x, y, z)$, where again (x, y, z) is any point of the flow, we would expect the random variable $\delta u(r)$ to be normal, in particular to have a skewness

$$\gamma(r) = E[\delta u(r)]^3 / \{E[\delta u(r)]^2\}^{3/2}$$

equal to zero and a flatness

$$\phi(r) = E[\delta u(r)]^4 / \{E[\delta u(r)]^2\}^2$$

equal to three, the appropriate values for a normal random variable.[1] What is actually found to be the case instead is that $\gamma(r)$ and $\phi(r)$ are appreciably different from zero and three, respectively, for values of r such that the correlation function

$$f(r) = Eu_1(x + r, y, z)u_1(x, y, z)/E[u_1(x, y, z)]^2$$

is appreciably different from zero. For experimental value of $\gamma(r)$ and $\phi(r)$ and a detailed discussion of the statistica structure of turbulence, we refer to the last chapter o Batchelor's monograph.[2]

The state of affairs just described raises the followin question: Is it possible by suitably choosing the sho rate and the pulse shape to construct shot noise whic mimics the statistical structure of turbulence, *i.e.*, whic is quite accurately univariate normal but exhibits marke departures from bivariate normality at close ranges? Th answer is in the affirmative, as we shall now show. Fo simplicity, we consider first the case of a one-dimensiona random process

$$u(t) = \sqrt{\frac{3}{\rho}} \sum_{i=-\infty}^{\infty} \beta_i s(t - t_i) - c\sqrt{\rho}, \qquad (1$$

where the t_i are the random occurrence times of the even in a stationary Poisson process with an average rate ρ events per second. The β_i are a family of independen random variables, all uniformly distributed over the un interval (0, 1). The centering constant c is chosen to b $\sqrt{3}/2 \int_{-\infty}^{\infty} s(t)\, dt$, which assures that $Eu(t) = 0$, an the normalization is such that $Eu^2(t)$ is independent of and is in fact equal to $\int_{-\infty}^{\infty} s^2(t)\, dt$. (A similar constructio has been given elsewhere.[3])

* Received by the PGIT, April 15, 1960. The research reported in this paper has been supported by the AF Cambridge Res. Center, Air Res. and Dev. Command, under Contract No. AF19(604)5238.

† Inst. of Mathematical Sciences, New York University in New York.

[1] The symbol E denotes the expectation value or ensemble average.

[2] G. K. Batchelor, "The Theory of Homogeneous Turbulence Cambridge University Press, Cambridge, Eng.; 1953.

[3] R. A. Silverman, "An isospectral family of random processe IRE Trans. on Information Theory, vol. IT-6, pp. 485–49 September, 1960.

We must now suitably choose the pulse shape $s(t)$ and the rate ρ so that the random variable $u(t)$ is quite accurately normal, while for small τ the difference random variable $u(t+\tau) - u(t)$, in particular $u'(t) \equiv du(t)/dt$ corresponding to the limit $\tau \to 0$, is quite markedly non-normal.[4] To achieve this, we choose $s(t)$ to be the function

$$s(t) = \begin{cases} t/2\epsilon, & 0 \le t < 2\epsilon, \\ 1, & 2\epsilon \le t < \alpha - \epsilon, \\ (\alpha - t)/\epsilon, & \alpha - \epsilon \le t < \alpha, \\ 0, & \alpha \le t, \end{cases} \tag{2}$$

with derivative

$$s'(t) = \begin{cases} 1/2\epsilon, & 0 \le t < 2\epsilon, \\ 0, & 2\epsilon \le t < \alpha - \epsilon, \\ -1/\epsilon, & \alpha - \epsilon \le t < \alpha, \\ 0, & \alpha \le t, \end{cases}$$

where $\epsilon \ll \alpha$ is a small parameter to be adjusted later. Our choice of $s(t)$ is motivated by the fact that its support (α) is much larger than the support (3ϵ) of its derivative; the significance of this will emerge presently. We have also arranged to give $u'(t)$ a negative skewness, in order to resemble the turbulent velocity field.

We now use the fact that the semi-invariants μ_n of $u(t)$ are given by the formula

$$\mu_n = \rho\left(\frac{3}{\rho}\right)^{n/2} E\beta^n \int_{-\infty}^{\infty} s^n(t)\, dt,$$

where β is any of the β_i.[5] In our case $E\beta^n = 1/(n+1)$, since β is uniformly distributed over the interval $(0, 1)$; the inclusion of the random variables β_i in (1) is to assure that the distributions of $u(t)$ and $u'(t)$ contain no delta function terms. In view of the relations $\mu_2 = m_2$, $\mu_3 = m_3$ and $\mu_4 = m_4 - 3m_2^2$ between the semi-invariants μ_2, μ_3, and μ_4 and the central moments m_2, m_3, and m_4, we find that the skewness of $u(t)$ is given by

$$\gamma = \frac{\mu_3}{\mu_2^{3/2}}$$

and its flatness by

$$\phi = 3 + \frac{\mu_4}{\mu_2^2}.$$

For the case of the pulse shape (2), we find that

$$\gamma = \frac{3}{4}\sqrt{\frac{3}{\rho}}\,\frac{\alpha - \frac{9}{4}\epsilon}{(\alpha - 2\epsilon)^{3/2}} \sim \frac{1.3}{\sqrt{\rho\alpha}}, \quad \text{if} \quad \epsilon \ll \alpha,$$

and

$$\phi = 3 + \frac{9}{5\rho}\,\frac{\alpha - \frac{12}{5}\epsilon}{(\alpha - 2\epsilon)^2} \sim 3 + \frac{1.8}{\rho\alpha}, \quad \text{if} \quad \epsilon \ll \alpha,$$

so that for small ϵ/α, the condition $\sqrt{\rho\alpha} \gg 1$ assures that $u(t)$ is quite accurately normal.[6] On the other hand, differentiating (1), we see that the univariate distribution of the process $u'(t)$ is governed by the semi-invariants

$$\mu_n' = \rho\left(\frac{3}{\rho}\right)^{n/2} E\beta^n \int_{-\infty}^{\infty} [s'(t)]^n\, dt,$$

with corresponding skewness

$$\gamma' = -\frac{3}{4}\,\frac{1}{\sqrt{2\rho\epsilon}} \sim -0.5\,\frac{1}{\sqrt{\rho\epsilon}},$$

and flatness

$$\phi' = 3 + \frac{0.9}{\rho\epsilon}.$$

Examining the expressions for γ, ϕ, γ', and ϕ', we see that by satisfying the conditions $\sqrt{\rho\alpha} \gg 1$ and $\sqrt{\rho\epsilon} \sim 1$, we can arrange to have simultaneously a quite accurately normal distribution of $u(t)$ and a markedly non-normal distribution of $u'(t)$. Moreover, the two conditions $\sqrt{\rho\alpha} \gg 1$ and $\sqrt{\rho\epsilon} \sim 1$ are compatible, provided only that $\sqrt{\alpha/\epsilon} \gg 1$. For example, if $\alpha = 1$, $\epsilon = 10^{-4}$ and $\rho = 10^4$, we have $\gamma \sim 0$, $\phi \sim 3$, $\gamma' \sim -0.5$, and $\phi' \sim 3.9$.

The skewness $\gamma(\tau)$ and the flatness $\phi(\tau)$ of the difference $u(t+\tau) - u(t)$ can be calculated in just the same way by observing (following a suggestion of Gilbert) that the process $u(t+\tau) - u(t)$ is given by (1) if we replace $s(t)$ by $s(t+\tau) - s(t)$. It is found that as $\tau \to 0$, $\gamma(\tau)$ and $\phi(\tau)$ reduce continuously to the limiting values γ' and ϕ', and that as τ approaches the correlation distance α of the process $u(t)$, $\gamma(\tau)$ and $\phi(\tau)$ approach the normal values of zero and three.[7] Thus, $u(t)$ resembles the turbulent velocity field by having a markedly non-normal bivariate distribution at close ranges.

Precisely the same kind of construction can be carried out in three dimensions by replacing the random times t_i by a spatial Poisson distribution and replacing the pulses $s(t)$ by three-dimensional "blobs." For example, suppose that

$$u(x, y, z) = \sqrt{\frac{3}{\rho_V}} \sum_i \beta_i s(x - x_i) s(y - y_i) s(z - z_i) - c\sqrt{\rho_V},$$

[4] The parameter t is fixed but arbitrary.

[5] S. O. Rice, "Mathematical analysis of random noise," reprinted in the collection "Selected Papers on Noise and Stochastic Processes," N. Wax, ed., Dover Publications, Inc., New York, N. Y., pp. 150–157; 1954.

[6] See the remarks in D. Middleton, "An Introduction to Statistical Communication Theory," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 505–506; 1960.

[7] The correlation function of $u(t)$ is

$$\int_{-\infty}^{\infty} s(t+\tau)s(t)\, dt \Big/ \int_{-\infty}^{\infty} s^2(t)\, dt,$$

which vanishes for $\tau \ge \alpha$. The function $\gamma(\tau)$ undergoes a sign change in the interval $(0, \alpha)$.

where the function s is the same as in (2), while this time, ρ_V is the *volume* density of the points (x_i, y_i, z_i) of a homogeneous spatial Poisson process and the summation is over all the "centers" (x_i, y_i, z_i). For each point (x_i, y_i, z_i), β_i is an independent random variable, uniformly distributed over the interval (0, 1). The centering constant c is now

$$\frac{\sqrt{3}}{2}\left\{\int_{-\infty}^{\infty} s(x)\, dx\right\}^3,$$

and

$$Eu(x, y, z) = 0, \qquad Eu^2(x, y, z) = \left\{\int_{-\infty}^{\infty} s^2(x)\, dx\right\}^3.$$

For the skewness γ and the flatness ϕ of the random variable $u(x, y, z)$, we now have

$$\gamma \sim \frac{1.3}{\sqrt{\rho_V \alpha^3}}, \qquad \phi \sim 3 + \frac{1.8}{\rho_V \alpha^3},$$

while for the skewness γ' and flatness ϕ' of the random variable $(\partial/\partial x)\, u(x, y, z)$, we find[8]

$$\gamma' \sim -\frac{3}{4}\frac{1}{\sqrt{2\rho_V \alpha^2 \epsilon}} \sim -0.5\frac{1}{\sqrt{\rho_V \alpha^2 \epsilon}},$$

$$\phi' \sim 3 + \frac{0.9}{\rho_V \alpha^2 \epsilon}.$$

If $\sqrt{\rho_V \alpha^3} \gg 1$ and $\sqrt{\rho_V \alpha^2 \epsilon} \sim 1$, the random field $u(x, y, z)$ is quite accurately univariate normal but exhibits marked departures from bivariate normality at close ranges. These two conditions are compatible, provided that $\sqrt{\alpha/\epsilon} \gg 1$, as before. For example, if $\alpha = 1$, $\epsilon = 10^{-}$ and $\rho_V = 10^4$, we have $\gamma \sim 0$, $\phi \sim 3$, $\gamma' \sim -0.5$ and $\phi' \sim 3.9$.

In conclusion, we see that when a large number of elementary waveforms are superimposed, even with high density and considerable overlap, there is no reason to expect *a priori*, in the absence of detailed information about the shape of the waveforms, that the resulting process is accurately normal.

[8] We use the obvious modifications of the formulas for the semi-invariants μ_n and μ_n'.

Maximum-Weight Group Codes for the Balanced *M*-Ary Channel*

CARL W. HELSTROM†

Summary—The construction of (n, k) group alphabets is discussed for the balanced M-ary channel, where M is the power of a prime. In this channel all M digits are equally likely to be in error, and an incorrect digit is equally likely to be any digit besides the one sent. The alphabets are formed by taking n columns of the modular representation table of the Abelian group of k-tuples of elements from the Galois field $GF(M)$ under digitwise addition. The formation and properties of that table are described. Attention is focused on alphabets in which all letters except the n-tuple of 0's have the maximum number of non-null elements. Tables of such alphabets are given for $M = 2$, $k = 2, 3, 4$; $M = 3$, $k = 2, 3$; and $M = 4$, $k = 2, 3$.

I. The Balanced *M*-Ary Channel

In the M-ary channel there are M different signals available, each corresponding to one of a set of M digits. One signal is sent at a time, say every T seconds. The channel is termed "balanced," if all digits are correctly received with the same probability, and if an incorrect digit is equally likely to be any digit other than the one transmitted.

An example of a balanced M-ary channel is one with M orthogonal signals $f_j(t)$ of duration T, $1 \leq j \leq M$:

$$\int_0^T f_i(t) f_j(t)\, dt = 0$$

for $i \neq j$. All signals are received with the same energy in white Gaussian noise of bandwidth much larger than $1/T$. The receiver contains M filters in parallel, each matched to one of the orthogonal signals. If the signals are narrow-band modulations of a carrier with random phase, the matched filters are followed by detectors whose outputs are measured at the end of each reception interval of duration T. The receiver emits that digit corresponding to the filter whose detected output is largest. Reiger has given a formula for the probability

* Received by the PGIT, March 17, 1960; revised manuscript received, May 19, 1960.

† Westinghouse Res. Labs., Pittsburgh, Pa.

that a digit is correctly received:[1]

$$q = \frac{1}{M} \sum_{r=1}^{M} (-1)^{r-1} \binom{M}{r} \exp\left[-(r-1)E/rN\right],$$

where N is the spectral density of the noise. A frequency-shift-keying system with M well-separated frequencies can be represented as a balanced M-ary channel.

Error-correction codes suitable for such a channel have been proposed by several writers, mostly for special values of M.[2-5] The codes studied by Lee[6] and in most of Ulrich's paper[2] are designed for a different kind of nonbinary channel, one whose M digits are somehow ordered, with the noise more likely to change a digit into its nearest neighbors than into other, more distant digits. Some properties of single-error correcting codes for the balanced M-ary channel, M the power of a prime, are described in another paper.[7]

II. Group Alphabets for the Balanced M-Ary Channel

The letters of a group alphabet can be considered as elements of an Abelian group.[8] In the group alphabets described here, the letters are n-tuples of the M-ary digits of the balanced channel, and M is the power of a prime. The total number of such n-tuples is M^n, $L = M^k$ of which are selected as code letters. We call such an alphabet an "(M, n, k) alphabet." Many of Slepian's results for binary (n, k) group alphabets can be applied with obvious changes to these (M, n, k) alphabets.

Each digit is assigned to an element of the Galois field of order M, abbreviated $GF(M)$.[5,9] The alphabet is then a subgroup of order M^k of the group C_n^M of all M^n n-tuples, in which the group operation is digitwise addition, performed by the rules for the Galois field.[10] We denote by 0 the unit element under addition and the digit assigned to it, and by 1 the unit element under multiplication and its assigned digit. The remaining $M - 2$ digits are called "nonunit" digits, and all digits except 0 are called "non-null." We define the "weight" of an n-tuple as the number of non-null digits it contains. All alphabets include the letter $I = (000 \cdots 0)$ of weight $w_0 = 0$.

These group alphabets are well suited to the balanced M-ary channel when messages are coded so that all letters are sent equally often. The received sets of n digits are to be decoded by assigning to each the code letter with the largest posterior probability of having been sent. This can be done by dividing the group C_n^M into cosets, using as the leader of each coset the element with the smallest weight, and subtracting[11] from the digits of each received set those of the leader of its coset.

One can think of the received n-tuple as the digitwise sum of the transmitted letter and an n-tuple caused by the noise. In order for a letter to be correctly received, the noise n-tuple must be either the letter $I = (000 \cdots 0)$ or one of the $(M^{n-k} - 1)$ coset leaders. Therefore, the probability Q of correctly receiving a code letter is

$$Q = q^n + \sum_i q^{n-w(i)} p^{w(i)},$$

$$p = (1 - q)/(M - 1),$$

$$1 \le i \le M^{n-k} - 1,$$

where $w(i)$ is the weight of the ith coset leader.

III. The Modular Representation Table

Like binary group alphabets,[8] these (M, n, k) alphabets can be conveniently constructed by selecting columns from the modular representation table (MRT) of the Abelian group C_k^M of order M^k, which is formed by all k-tuples of elements of $GF(M)$ under digitwise field addition. This MRT contains M^k rows and columns, each labeled with a k-tuple of field elements from $GF(M)$. At the intersection of the row labeled $(b_1, b_2, \cdots b_k)$ and the column labeled $(a_1, a_2, \cdots a_k$ is found the field element

$$a_1b_1 + a_2b_2 + \cdots a_kb_k,$$

in which the addition and multiplication are performed according to the rules for the Galois field. The MRT for $M = 2$, $k = 4$ is given by Slepian.[8]

The columns of the MRT form a representation of the group C_k^M in the following sense. If the elements of any two columns are added row-by-row by the rules for the field, a new column results that is the same as the one labeled with the group product (digitwise sum) of the labels of the first two columns, because the field operations obey the distributive laws.[12] Henceforth, we omit the column and row of all '0's in the MRT.

The remaining columns and rows of the MRT fall into sets of $(M - 1)$ columns (or rows) that differ only by a multiplicative factor. By selecting one column and row of each set, we can write down a "reduced MRT" (RMRT) that has only

$$K = (M^k - 1)/(M - 1)$$

[1] S. Reiger, "Error rates in data transmission," Proc. IRE, vol. 46, p. 919; May, 1958.

[2] W. Ulrich, "Nonbinary error correction codes," *Bell Sys. Tech. J.*, vol. 36, pp. 1341–1388; November, 1957. See Sec. IV, pp. 1360–1367.

[3] M. J. E. Golay, "Notes on the penny-weighing problem, lossless symbol coding with nonprimes, etc.," IRE Trans. on Information Theory, vol. IT-4, pp. 103–109; September, 1958.

[4] H. S. Shapiro and D. L. Slotnick, "On the mathematical theory of error-correcting codes," *IBM J. Res. Dev.*, vol. 3, pp. 25–34; January, 1959.

[5] J. Cocke, "Lossless symbol coding with nonprimes," IRE Trans. on Information Theory, vol. IT-5, pp. 33–34; March, 1959.

[6] C. Y. Lee, "Some properties of nonbinary error-correcting codes," IRE Trans. on Information Theory, vol. IT-4, pp. 77–82; June, 1958.

[7] C. W. Helstrom, "Single-error correcting codes for nonbinary balanced channels," IRE Transactions on Information Theory, to be published.

[8] D. Slepian, "A class of binary signaling alphabets," *Bell Sys. Tech. J.*, vol. 35, pp. 203–234; January, 1956.

[9] B. M. Dwork and R. M. Heller, "Results of a geometric approach to the theory and construction of nonbinary, multiple error and failure correcting codes," 1959 IRE National Convention Record, pt. 4, pp. 123–129.

[10] B. L. van der Waerden, "Modern Algebra," F. Ungar Publishing Co., New York, N. Y., Sec. 37, pp. 115–119; 1949.

[11] "Subtraction" is the operation inverse to Galois field addition.

[12] van der Waerden, *op. cit.*, p. 32.

rows and columns; it exhibits well enough the structure of the original MRT. (For $M = 2$, the RMRT is the same as the MRT.) In Table I, we give an RMRT for $M = 4$, $k = 3$; it is a 21×21 matrix. The field elements are now 0, 1, α, and β; they obey the rules[9]

$$x + x = 0, \qquad x \cdot 1 = x, \qquad (x \quad \text{any element}) \qquad \beta = \alpha \cdot \alpha, \qquad 1 + \alpha + \beta = 0, \qquad \alpha \cdot \beta = \beta \cdot \alpha = 1. \tag{1}$$

TABLE I

REDUCED MODULAR REPRESENTATION TABLE OF C_3^4

	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	0	1	1	1	0	0	0	1	α	β	1	1	1	α	α	α	β	β	β
	1	0	0	1	α	β	1	α	β	0	0	0	1	α	β	1	α	β	1	α	β
001	1	0	0	1	α	β	1	α	β	0	0	0	1	α	β	1	α	β	1	α	β
010	0	1	0	1	1	1	0	0	0	1	α	β	1	1	1	α	α	α	β	β	β
100	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
011	1	1	0	0	β	α	1	α	β	1	α	β	0	β	α	β	0	1	α	1	0
01α	α	1	0	β	α	0	α	β	1	1	α	β	β	α	0	0	1	β	1	0	α
01β	β	1	0	α	0	β	β	1	α	1	α	β	α	0	β	1	β	0	0	α	1
101	1	0	1	1	α	β	0	β	α	1	1	1	0	β	α	0	β	α	0	β	α
10α	α	0	1	α	β	1	β	α	0	1	1	1	β	α	0	β	α	0	β	α	0
10β	β	0	1	β	1	α	α	0	β	1	1	1	α	0	β	α	0	β	α	0	β
110	0	1	1	1	1	1	1	1	1	0	β	α	0	0	0	β	β	β	α	α	α
1α0	0	α	1	α	α	α	1	1	1	β	α	0	β	β	β	α	α	α	0	0	0
1β0	0	β	1	β	β	β	1	1	1	α	0	β	α	α	α	0	0	0	β	β	β
111	1	1	1	0	β	α	0	β	α	0	β	α	1	α	β	α	1	0	β	0	1
11α	α	1	1	β	α	0	β	α	0	0	β	α	α	β	1	1	0	α	0	1	β
11β	β	1	1	α	0	β	α	0	β	0	β	α	β	1	α	0	α	1	1	β	0
1α1	1	α	1	β	0	1	0	β	α	β	α	0	α	1	0	β	0	1	1	α	β
1αα	α	α	1	0	1	β	β	α	0	β	α	0	1	0	α	0	1	β	α	β	1
1αβ	β	α	1	1	β	0	α	0	β	β	α	0	0	α	1	1	β	0	β	1	α
1β1	1	β	1	α	1	0	0	β	α	α	0	β	β	0	1	1	α	β	α	1	0
1βα	α	β	1	1	0	α	β	α	0	α	0	β	0	1	β	α	β	1	1	0	α
1ββ	β	β	1	0	α	1	α	0	β	α	0	β	1	β	0	β	1	α	0	α	1

The MRT and the RMRT can be subdivided into blocks by taking together all columns (and rows) with given numbers of '0's in their labels. In Table I the blocks are indicated by heavy lines. We denote by H_m the set of k-tuples with m non-null elements, as well as the sets of columns and rows labeled by those k-tuples. In a block with rows in H_m and columns in H_n, each row has the same number μ_{mn} of '0's. In the Appendix we derive a formula for the numbers μ_{mn}, which enables one to count these '0's and learn something of the structure of the MRT for C_k^M without writing out the whole table.

In particular, the number of '0's in a column of the MRT is M^{k-1}. Therefore, in an (M, n, k) alphabet the sum of the weights w_i of all the letters is

$$\sum_{k=0}^{L} w_k = n(M^k - M^{k-1}) = nM^{k-1}(M - 1) = S, \tag{2}$$

and the nonzero weights form a partition of this number S into $L - 1$ parts (leaving out $w_0 = 0$). (For $M = 2$ this is Slepian's Proposition 6.[8])

IV. MAXIMUM-WEIGHT (M, n, k) ALPHABETS

The distance between two code letters can be conveniently defined as the number of places in which they differ. The greater this distance, the less likely one lett is to be received when the other is sent. Now this distanc is the weight of a third letter of the alphabet, found b subtracting the two letters digit by digit. Therefore, th (M, n, k) alphabet, constructed so that the weights w of all letters other than $I = (000 \cdots 0)$ are as large a possible, can be expected to have nearly the smalles attainable probability of error. We call such alphabet "maximum-weight alphabets."

If instead of quantizing each received signal into on of the M digits, the receiver bases its decisions about th transmitted letters on the amplitudes of the n-associate received signals, using the principle of maximum likel hood, the maximum-weight alphabet yields the larges possible probability of correct reception among all $(M, n, k$ alphabets, at least in the limit of large signal-to-nois ratio. When the received signals are first quantized int digits, n-tuples of which are then decoded into letter of the alphabet, the maximum-weight alphabet does no always yield the minimum probability of error, as Slepia has shown for the (2, 7, 3) code.[8] One expects, howeve that the maximum-weight alphabet will be less vulnerabl to noise than most of the large number of possible $(M, n, k$ alphabets.

MacDonald,[13] and Bose and Kuebler[14] have discusse the construction of binary group codes of maximu weight. Bose and Kuebler treat the problem in terms o a finite projective geometry, and their approach coul no doubt be extended to M-ary alphabets when M i the power of a prime. The first k elements of the K colum of the RMRT are the homogeneous coordinates of th points in such a finite geometry.

We define a "maximal partition" of $S = nM^{k-1}(M -$ 1 into $L - 1$ parts as one in which all the parts are a large as possible, without regard to whether a group co with those weights exists—usually none does. If w_1 the smallest nonzero weight, it satisfies the inequalit

$$w_1 \leq S/(L - 1), \qquad L = M^k,$$

and if w_2 is the next smallest ($w_2 \geq w_1$),

$$w_2 \leq (S - w_1)/(L - 2),$$

and so forth. The maximal partition is easily calculat by taking the largest integers permitted by these i equalities.

If the letter length n is a multiple of K:

$$n = Ks, \qquad K = (M^k - 1)/(M - 1), \qquad ($$

the maximal partition is one in which all weights are equa

$$w_j = sM^{k-1}, \qquad 1 \leq j \leq M^k - 1. \qquad ($$

For $s = 1$, alphabets having these weights can be form

[13] J. E. MacDonald, "Design methods for maximum minimu distance error-correcting codes," *IBM J. Res. Dev.*, vol. 4, p 43–57; January, 1960.

[14] R. C. Bose and R. R. Kuebler, Jr., "A geometry of bina sequences associated with group alphabets in information theory *Ann. Math. Stat.*, vol. 31, pp. 113–139; March, 1960.

by using as letters the K rows of the RMRT, along with the digitwise product of each row with the $(M - 2)$ nonunit field elements Stated otherwise, one picks one column of the MRT from each of the K sets of columns differing only by a multiplicative factor. For $s > 1$, one repeats each of these letters s times

When the length n is not a multiple of K:

$$n = Ks + h, \qquad s \geq 0, \quad 1 \leq h < K, \tag{5}$$

the maximum-weight codes are formed by appending certain columns of the RMRT to the s-times repeated columns used for $n = Ks$ These h extra columns are to be chosen so that the weights of the rows of an $h \times K$ matrix formed from them are all as large as possible The h columns to be used do not depend on s, and we shall take $s = 0$. As before, the list is to be expanded to $(L - 1)$ letters by multiplying each row by the $(M - 2)$ nonunit field elements to obtain the complete alphabet. There are many equivalent codes with the same sets of weights, for the best choices of columns are not unique. Each column, for instance, can be multiplied through by a nonunit field element to yield a new column that serves as well. Still other transformations and permutations of columns are possible that yield equivalent codes.

We conjecture, but have not proved, that in the set of h extra columns for a group code of maximum weight, no column will appear twice There are exactly enough columns in the RMRT for this to be possible for all values of h. Plausibility is lent to the conjecture by the observation that one can usually pass from $h = h'$ to $h = h' + 1$ by adjoining a previously unused column of the RMRT to the optimum set for $h = h'$, choosing the new one so that it has non-null elements in as many of the rows of minimum weight for $h = h'$ as possible. However, for certain values of k and $h = h'$ it has been found necessary to drop a column and add two new ones in passing to $h = h' + 1$. For still larger values of k, more extensive revisions may be needed at certain values of h.

In Table II we list, using Slepian's notation,[8] a set of choices of extra columns yielding maximum-weight codes, and the resulting letter weights, for $M = 2$, $k = 2, 3, 4$. For instance, when $k = 4$, (124) stands for the binary number (1101), and "3^5" means that there are five letters of weight 3 in the alphabet. In Table III we give the extra columns and weights for $M = 3$, $k = 2, 3$, and in Table IV we list the same for $M = 4$. For $M = 3$ the digits are the integers modulo 3: 0, 1, 2; for $M = 4$ we use 0, 1, α, and β as in (1).

We shall now briefly discuss the procedure for finding these maximum-weight codes. For values of h in $1 \leq h \leq k$, one picks any h different columns from the part of the RMRT labeled by H_1. As h increases from k, one adds columns from H_k until these are exhausted, or until too great an imbalance among the weights seems to occur. As for $M = 2$, $k = 4$, $h = 6$, it may then be necessary to drop some columns and add new columns from other parts of the RMRT to find the best selection.

For values of h near K, one proceeds by successively dropping columns from the complete RMRT. The best procedure here seems to be to choose two elements of C_k^M not in H_1 and not differing by a multiplicative factor, and to form the subgroup they generate. One first drops those columns of the RMRT labeled with elements of that subgroup. One then takes an element not in the subgroup and forms the subgroup generated by it and the elements of the preceding subgroup. The next columns to be dropped from the RMRT are those labeled by the elements of this new subgroup, but not already removed. One continues thus until the largest subgroup has been formed and the associated columns removed from the RMRT. With possibly a few modifications this procedure should yield maximum-weight codes for $M^{k-1} \leq h < K$.

Appendix

Distribution of '0's in the Modular Representation Table

Our problem is to find the number μ_{mn} of zeros in each row of a block of the modular representation table with row labels in H_m and column labels in H_n, that is, the number of ways of obtaining

$$a_1b_1 + a_2b_2 + \cdots a_kb_k = 0$$

when n of the a_i's and m of the b_i's differ from '0'. Since μ_{mn} is the same for all rows in the block, we choose

$$b_1 = b_2 = \cdots b_m = 1, \qquad b_{m+1} = \cdots = b_k = 0.$$

Then μ_{mn} is the number of ways $(a_1 + a_2 + \cdots a_m)$ equals 0 when $(a_1, a_2, \cdots a_k)$ is in H_n.

We fix at r the number of non-null a_i's among the first m, with $n - r$ remaining in the last $k - m$ places. Later we sum over r. There are

$$\binom{m}{r}(M-1)^r\binom{k-m}{n-r}(M-1)^{n-r}$$

ways of choosing the a_i's to satisfy this condition; of these, a fraction $\tau_r/(M - 1)^r$ yield $a_1 + a_2 + \cdots a_m = 0$. Then $(\tau_0 = 1)$

$$\mu_{mn} = \sum_{r=0}^{m} \tau_r \binom{m}{r}\binom{k-m}{n-r}(M-1)^{n-r}.$$

We must now find τ_r, which is just the number of ways

$$s_r = a_1 + a_2 + \cdots a_r$$

can vanish when the r a_i's, $1 \leq i \leq r$, run through the $(M - 1)$ non-null field elements independently. Let τ_r' be the number of ways of getting

$$s_r = a_1 + a_2 + \cdots a_r = x \neq 0.$$

TABLE II

MAXIMUM-WEIGHT GROUP CODES ($M = 2$)

	h	Labels of Extra Columns from MRT	Weights w_i	Maximal Partition
$k = 2$	1	3	$0\ 1^2$	$0\ 1^2$
	2	1, 2	$1^2\ 2$	$1^2\ 2$
$k = 3$	1	1	$0^3\ 1^4$	$0^3\ 1^4$
	2	1, 2	$0\ 1^4\ 2^2$	$1^6\ 2$
	3	1, 2, 3	$1^3\ 2^3\ 3$	$1^2\ 2^5$
	4	1, 2, 3, 123	$2^6\ 4$	$2^5\ 3^2$
	5	1, 2, 3, 12, 13	$2^2\ 3^4\ 4$	$2\ 3^6$
	6	1, 2, 3, 12, 13, 23	$3^4\ 4^3$	$3^4\ 4^3$
$k = 4$	1	1	$0^7\ 1^8$	$0^7\ 1^8$
	2	1, 2	$0^3\ 1^8\ 2^4$	$1^{14} 2$
	3	1, 2, 3	$0\ 1^6 2^6 3^2$	$1^6\ 2^9$
	4	1, 2, 3, 4	$1^4 2^6 3^4 4$	$2^{13} 3^2$
	5	1, 2, 3, 4, 1234	$2^{10} 4^5$	$2^5\ 3^{10}$
	6	1, 2, 3, 4, 134, 234	$2^3\ 3^8\ 4^3 6$	$3^{12} 4^3$
	7	1, 2, 3, 4, 124, 134, 234	$3^7\ 4^7\ 7$	$3^4\ 4^{11}$
	8	1, 2, 3, 4, 124, 134, 234 + 123	$4^{14} 8$	$4^{11} 5^4$
	9	1, 2, 3, 4, 124, 134, 234 + 13	$4^6\ 5^8\ 8$	$4^3\ 5^{12}$
	10	1, 2, 3, 4, 124, 134, 234 + 14	$4^2 5^8 6^4 8$	$5^{10} 6^5$
	11	1, 2, 3, 4, 124, 134, 234 + 23	$5^6 6^6 7^2 8$	$5^2\ 6^{13}$
	12	1, 2, 3, 4, 124, 134, 234 + 24	$6^{12} 8^3$	$6^9\ 7^6$
	13	1, 2, 3, 4, 124, 134, 234 + 1234	$6^4\ 7^8\ 8^3$	$6\ 7^{14}$
	14	1, 2, 3, 4, 124, 134, 234 + 34	$7^8\ 8^7$	$7^8\ 8^7$

TABLE III

MAXIMUM-WEIGHT GROUP CODES ($M = 3$)

	h	Labels of Extra Columns from MRT	Weights w_i
$k = 2$	1	01	$0^2 1^6$
	2	01 + 10	$1^4 2^4$
	3	01 + 11	$2^6 3^2$
	4	01 + 12	3^8
$k = 3$	1	001	$0^8\ 1^{18}$
	2	001 + 010	$0^2\ 1^{12}\ 2^8$
	3	001 + 100	$1^6\ 2^{12}\ 3^{12}$
	4	001 + 111	$2^{12}\ 3^8\ 4^6$
	5	001 + 112	$2^2\ 3^{14}\ 4^6\ 5$
	6	001 + 120	$3^4\ 4^{18}\ 6^4$
	7	001 + 121	$4^{10}\ 5^{12}\ 6^2$
	8	001 + 011	$5^{16}\ 6^8\ 8^2$
	9	001 + 102	$6^{24}\ 9^2$
	10	001 + 122	$6^6\ 7^{18}\ 9^2$
	11	001 + 012	$7^{12}\ 8^{12}\ 9^2$
	12	001 + 101	$8^{18}\ 9^8$
	13	001 + 110	9^{26}

TABLE IV

MAXIMUM-WEIGHT GROUP CODES ($M = 4$)

	h	Labels of Extra Columns from MRT	Weights w_i
$k = 2$	1	01	$0^3\ 1^{12}$
	2	01 + 10	$1^6\ 2^9$
	3	01 + 11	$2^9\ 3^6$
	4	01 + 1α	$3^{12}\ 4^3$
	5	01 + 1β	4^{15}
$k = 3$	1	001	$0^{15}\ 1^{48}$
	2	001 + 010	$0^3\ 1^{24}\ 2^{36}$
	3	001 + 100	$1^9\ 2^{27}\ 3^{27}$
	4	001 + 111	$2^{18}\ 3^{24}\ 4^{21}$
	5	001 + $1\beta\alpha$	$3^{30}\ 4^{15}\ 5^{18}$
	6	001 + $1\alpha\beta$	$4^{45}\ 6^{18}$
	7	001 + $1\beta\beta$	$4^9\ 5^{36}\ 6^6\ 7^{12}$
	8	001 + 01α	$5^{18}\ 6^{30}\ 7^6\ 8^9$
	9	001 + 101	$6^{27}\ 7^{27}\ 9^9$
	10	001 + 110	$6^3\ 7^{33}\ 8^{18}\ 9^3\ 10^6$
	11	001 + $1\alpha\alpha$	$7^6\ 8^{39}\ 9^{12}\ 11^6$
	12	001 + 01β	$8^9\ 9^{48}\ 12^6$
	13	001 + 11α	$9^{21}\ 10^{36}\ 12^3\ 13^3$
	14	001 + $1\alpha 1$	$10^{33}\ 11^{24}\ 12^3\ 14^3$
	15	001 + 10β	$11^{45}\ 12^{15}\ 15^3$
	16	001 + $1\beta 0$	$12^{60}\ 16^3$
	17	001 + 011	$12^{12}\ 13^{48}\ 16^3$
	18	001 + 10α	$13^{24}\ 14^{36}\ 16^3$
	19	001 + $1\alpha 0$	$14^{36}\ 15^{24}\ 16^3$
	20	001 + 11β	$15^{48}\ 16^{15}$
	21	001 + $1\beta 1$	16^{21}

Because of the symmetry among the non-null group elements, τ_r' is independent of x. The following recursion relations connect τ_r and τ_r':

$$\tau_{r+1} = (M - 1)\tau_r'$$

$$\tau_{r+1}' = \tau_r + (M - 2)\tau_r'.$$

The former of these indicates that one can get $s_{r+1} =$ in $(M - 1)$ ways, namely by taking $a_{r+1} = -x$ wh $s_r = x$, $x \neq 0$; $s_r = x$ was formed in τ_r' ways. Similarl one gets $s_{r+1} = x \neq 0$ by adding $a_{r+1} = x$ to $s_r =$ or by adding the element $x - y$ to $s_r = y \neq 0$. The form can be done in one way, the latter in $(M - 2)$ ways.

The solution of the above difference equations with itial conditions $\tau_1 = 0$, $\tau_1' = 1$ is

$$\tau_r = (M-1)[(M-1)^{r-1} - (-1)^{r-1}]/M$$

$$\tau_r' = [(M-1)^r - (-1)^r]/M.$$

hus we finally get

$$_{nn} = M^{-1}\left\{\binom{k}{n}(M-1)^n + \sum_{r=0}^{m}\binom{m}{r}\binom{k-m}{n-r}(-1)^r(M-1)^{n-r+1}\right\}.$$

It is often convenient to use the generating functions

$$_m(z) = \sum_{n=0}^{k} \mu_{mn} z^n = M^{-1}\{[1+(M-1)z]^k + (M-1)[1+(M-1)z]^{k-m}(1-z)^m\}.$$

particular, the total number of '0's in a column or w of the MRT is

$$F_m(1) = \sum_{n=0}^{k} \mu_{mn} = M^k, \qquad m = 0,$$
$$= M^{k-1}, \qquad m \neq 0.$$

articularly simple values of μ_{mn} are:

$$\mu_{m0} = 1,$$
$$\mu_{mi} = (M-1)(k-m),$$
$$\mu_{mk} = M^{-1}(M-1)^{k-m+1}[(M-1)^{m-1} + (-1)^m].$$

The symmetry relation

$$\binom{k}{m}\mu_{mn}(M-1)^m = \binom{k}{n}\mu_{nm}(M-1)^n$$

follows from the symmetry of the MRT. To find the numbers of '0's in the blocks of the reduced MRT, divide all the μ_{mn}'s by $(M-1)$, $1 \leq (m, n) \leq k$.

For a binary MRT ($M = 2$), one has the simpler generating function

$$F_m(z) = \tfrac{1}{2}[(1+z)^k + (1+z)^{k-m}(1-z)^m].$$

For $k = 5$, $M = 2$, for instance, the matrix $\|\mu_{mn}\|$ is

$$\|\mu_{mn}\| = \begin{bmatrix} 1 & 5 & 10 & 10 & 5 & 1 \\ 1 & 4 & 6 & 4 & 1 & 0 \\ 1 & 3 & 4 & 4 & 3 & 1 \\ 1 & 2 & 4 & 6 & 3 & 0 \\ 1 & 1 & 6 & 6 & 1 & 1 \\ 1 & 0 & 10 & 0 & 5 & 0 \end{bmatrix}.$$

ACKNOWLEDGMENT

The author is indebted to E. Scherer, Dr. E. Muehldorf, and Dr. R. Filipowsky of Advanced Development, Westinghouse Electronics Division, Baltimore, Md., for discussions which stimulated the work reported here.

CORRECTION

The Editor wishes to call attention to the following corrections in G. R. Welti's paper "Quaternary Codes for Pulsed Radar," which appeared on pages 400–408 of the June, 1960, issue of these TRANSACTIONS.

1) Page 402, column 2, next to last equation: Right-hand side should read

$$\begin{bmatrix} \alpha & \alpha \\ \alpha & \beta \end{bmatrix}.$$

2) Page 405, column 1, four and one-half inches from bottom of page: The equation should begin with

$$m_p = \begin{cases} \nu \sum \cdots \\ \nu \sum \cdots \end{cases}.$$

3) Page 408, column 2, second equation: Left-hand side should read

$$D_i^k + \underline{D}_i^k + A_1^k + A_k^k.$$

Correspondence

Second-Order Properties of the Pre-Envelope and Envelope Processes*

In a recent paper, Dugundji[1] introduced a generalized definition of envelope based on the notion of "analytic signal" or "pre-envelope." Section IV of Dugundji's paper contains the results: 1) The time autocorrelation function of the Hilbert transform of a waveform is equal to the time autocorrelation of the original waveform, 2) The time cross-correlation function between a waveform and its Hilbert transform is the Hilbert transform of the time autocorrelation of the original waveform, 3) The time autocorrelation of the pre-envelope of a waveform is twice the pre-envelope of the time autocorrelation function of the original waveform. For ergodic processes, the same results remain true, indeed, when time averages are replaced by ensemble averages. The purpose of this note is to show that the results mentioned above of Section IV in Dugundji's paper remain true when *time averages* are replaced by *ensemble averages*, regardless of ergodicity, the only requirement being that the process be wide sense stationary. It will also be shown that a corollary of result 3), recently pointed out by Brown,[2,3] *the envelope of an ergodic process has a variance which is twice the variance of the original process* also remains true when the word *ergodic* is replaced by the words *wide sense stationary.*

The key to the proofs given here is to consider the frequency domain representation of the Hilbert transform. This approach simplifies the proofs since the frequency domain representation of the Hilbert transform is a simple all pass, nonrealizable "filter" or transfer function.

Letting $\{u(t)\}$ be a wide sense stationary process and $S_u(f)$ (with $\int_{-\infty}^{\infty} S_u(f)df < \infty$) be the spectral density of the process, the transfer function $Y(f)$ defines a linear operation on $\{u(t)\}$ if and only if[4]

$$\int_{-\infty}^{\infty} |\, Y(f)\,|^2\, S_u(f)\, df < \infty;$$

if $\{\hat{u}(t)\}$ is the result of operating with $Y(f)$ on $\{u(t)\}$ then $\{\hat{u}(t)\}$ is also wide sense stationary with the spectral density:

$$S_{\hat{u}}(f) = |\, Y(f)\,|^2\, S_u(f)$$

and the cross spectral density $S_{u,\hat{u}}(f)$ is given by

$$S_{u,\hat{u}}(f) = Y(f)S_u(f).$$

The linear operation induced by

$$Y(f) = \begin{cases} -i \;\; f > 0, \\ \;\;\,0 \;\; f = 0, \; i = \sqrt{-1} \\ \;\;\,i \;\; f < 0, \end{cases} \quad (1)$$

$$= -i\cdot \mathrm{sgn}(f)$$

will now be defined as the Hilbert transform, the familiar time domain representation of the Hilbert transform will be derived later from (1).

1) The spectral density of $\{\hat{u}(t)\}$, where $\{\hat{u}(t)\}$ is the Hilbert transform of $\{u(t)\}$, is, therefore, given by

$$S_{\hat{u}}(f) = |\, Y(f)\,|^2\, S_u(f)$$

$$= S_u(f), \quad \text{for} \quad f \neq 0$$

and therefore, the ensemble autocorrelation function of the $\{\hat{u}(t)\}$ process satisfies

$$R_{\hat{u}}(\tau) = R_u(\tau).$$

2) The cross spectral density between u and $\hat{u}$ is given by

$$S_{u,\hat{u}}(f) = Y(f)S_u(f)$$

$$= -i\cdot \mathrm{sgn}\,(f) S_u(f) \quad (2)$$

the ensemble cross-correlation $R_{u,\hat{u}}(\tau)$ is, therefore,

$$R_{u,\hat{u}}(\tau) = \int_{-\infty}^{\infty} Y(f)S_u(f)e^{i2\pi f}\,\tau df.$$

Since $|\, Y(f)\,| = 1$ and $S_u(f)$ is integrable $(-\infty, \infty)$, $R_{u,\hat{u}}(\tau)$ is defined for all τ, is bounded and continuous, and tends to zero as τ tends to $\pm\infty$.

As the Hilbert transform was defined in (1) in the frequency domain only, we have to show that in the time domain the relation between R_u and $R_{u,\hat{u}}$ is given by

$$R_{u,\hat{u}}(\tau) = \hat{R}_u(\tau) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{R_u(t)}{\tau - t}\, dt$$

where P denotes the principal value. In order to derive the last equation, let

$$Y(f, \delta, A) = \frac{1}{\pi}\left\{\int_{-A}^{-\delta} \frac{e^{-i2\pi ft}\, dt}{t} + \int_{\delta}^{A} \frac{e^{-i2\pi ft}\, dt}{t}\right\}, \qquad (A > \delta > 0)$$

$$= -\frac{2i}{\pi}\cdot(Si(2\pi fA) - Si(2\pi f\,\delta))$$

where

$$Si(x) = \int_0^x \frac{\sin y}{y}\, dy;$$

then $Y(f, \delta, A)$ is uniformly bounded in f, A, δ and $Y(f, \delta, A)$ tends to $Y(f)$ as $A \to \infty$ and $\delta \to 0$. Therefore,

$$\hat{R}(\tau) = \lim_{\substack{A\to\infty \\ \delta\to 0}} \int_{-\infty}^{\infty} Y(f, \delta, A) S_u(f) e^{i2\pi f}\,\tau df$$

Hence, by the convolution theorem,

$$\hat{R}(\tau) = \lim_{\substack{A\to\infty \\ \delta\to 0}} \left(\frac{1}{\pi}\int_{-A+\tau}^{-\delta+\tau} \frac{R_u(t)}{\tau - t}\, dt + \frac{1}{\pi}\int_{\delta+\tau}^{A+\tau} \frac{R_u(t)}{-t}\, dt\right)$$

$$= \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{R_u(t)}{\tau - t}\, dt. \quad (3)$$

3) The pre-envelope process $\{u(t) + i\hat{u}(t)\}$ is obtained from $\{u(t)\}$ by a linear operation with the transfer function $[1 + iY(f)]$, where $Y(f)$ is as defined in (1); hence,

$$S_{(u+i\hat{u})}(f) = |\, 1 + iY(f)\,|^2 \cdot S_u(f)$$

$$= 2[1 + iY(f)]S_u(f)$$

$$= \begin{cases} 4S_u(f), & f > 0 \\ 2S_u(f), & f = 0 \\ \;\;0, & f < 0, \end{cases} \quad (4)$$

and from (1), (2), (4), the ensemble auto-correlation function of the process $\{u(t) + i\hat{u}(t)\}$ is given by

$$R_{u+i\hat{u}}(\tau) = 2[R_u(\tau) + i\hat{R}_u(\tau)].$$

4) The envelope $V(t)$ of $u(t)$ is defined as $V(t) = [u^2(t) + \hat{u}^2(t)]^{1/2}$. Therefore the expectation of $V^2(t)$ is given by

$$E[V^2(t)] = R_u(0) + R_{\hat{u}}(0) = 2R_u(0).$$

Hence, the ensemble average of $V^2(t)$ is twice the ensemble average of $u^2(t)$.

As for the time domain representation of the operation induced by $Y(f)$ in (1) with respect to the process $\{u(t)\}$, using the same approximation as in the derivation of (3), it follows that $\hat{u}(t)$ is the limit in the mean of

$$\frac{1}{\pi}\int_{t-A}^{t-\delta} \frac{u(\tau)}{t - \tau}\, d\tau + \frac{1}{\pi}\int_{t+\delta}^{t+A} \frac{u(\tau)}{t - \tau}\, d\tau$$

as $A \to \infty$, $\delta \to 0$.

The representation of the Hilbert transform as a (all-pass, nonrealizable) "filter" or "transfer function" was used in this note to show that some known properties of the pre-envelope and envelope of ergodic processes are actually second-order properties

* Received by the PGIT, March 4, 1960; revised manuscript received, July 11, 1960. This work was supported by the Sci. Dept., Ministry of Defence, Israel.

[1] J. Dugundji, "Envelopes and pre-envelopes of real waveforms," IRE Trans. on Information Theory, vol. IT-4, pp. 53–57; March, 1958.

[2] W. M. Brown, "Some results on noise through circuits," IRE Trans. on Circuit Theory, vol. CT-6, pp. 217–227; May, 1959.

[3] J. L. Brown, Jr., "A property of the generalized envelope," IRE Trans. on Circuit Theory, vol. CT-6, p. 325; September, 1959.

[4] J. L. Doob, "Stochastic Processes," John Wiley and Sons, Inc., New York, N. Y., p. 534; 1953.

i.e., they hold for any wide sense stationary process. This representation can also be used to simplify the derivation of some other results on Hilbert transforms and pre-envelopes. The frequency domain analysis has an additional advantage over time domain analysis in this case, since it avoids manipulations with improper integrals.

Moshe Zakai
Sci. Dept.
Ministry of Defense
Israel.

Upper Bounds for Error Detecting and Correcting Codes*

The excellent paper submitted by N. Wax,[1] prompts me to contribute some additional thoughts for possible extension of the work on this subject.

From an examination of Table I, "Collected Results," in the original paper, the Hamming bound is apparently superior for low redundancy, or, equivalently, for small e/n ratios, whereas the improved "soft sphere" model is better for a range of moderate error correction.

This letter describes an approach to an error-correcting bound that can be made smaller than either of the aforementioned, under the condition that the number of correctable errors is large for a given n. Although a rigorous proof is not included in this letter, a logical development is offered.

The described bound, which may be referred to as the maximal codistant bound, arises from an argument that is in contrast to the usual "sphere" geometrical model; *i.e.*, Hamming[2] packs v disjoint spheres of error-correcting radius e to produce a minimum n space—of

$$v \sum_{i=0}^{e} \binom{n}{i} \leq 2^n.$$

On the other hand, Wax's bounds use clever geometrical techniques, which, as an analog, use rigid and elastic spheres in an attempt to fill the n space, and thereby arrive at a better estimate of the information density.

TABLE I

Collected Results

	R^2	0.75	1.25	1.75	2.25	2.75	3.25
n	e	1	3	2	4	5	6
7	v_1	74.2	12.4	4.1			
	v_H	16	4.4	2			
	v_2	18.2*	3.1	2.1			
	v_c	—	3.33	2			
	N_B	16	2	2			
8	v_1	199.3	25.6	7.1			
	v_H	28.4	6.9	2.8			
	v_2	20.3*	5.3	3			
	v_c	—	5	2.34			
	N_B	20**	4	2			
9	v_1	56.7	56.2	12.9	4.8		
	v_H	51.2	11.1	3.9	2		
	v_2	39.7*	9.2	4.2	2.7		
	v_c	—	10	2.8	2		
	N_B	38**	6	2	2		
10	v_2	1.69×10^3	128.7	25.0	7.6		
	v_H	93.1	18.3	5.8	2.7		
	v_2	82.2	16.6	6.1	3.4		
	v_c	—	—	3.5	2.25		
	N_B	68**	12	2	2		
11	v_1	5.29×10^3	312.5	50.6	13.5	5.5	
	v_H	170.7	30.6	8.8	3.6	2	
	v_2	154.8*	26.5*	8.6*	4.3	2.8	
	v_c	—	—	4.67	2.58	2	
	N_B	128	24	4	2	2	
12	v_1	1.72×10^4	781.2	106.9	25.5	8.8	
	v_H	315.1	51.8	13.7	5.2	2.6	
	v_2	346.8*	46.7*	12.7*	5.7	3.3	
	v_c	—	—	7	3	2.2	
	N_B	256	24	4	2	2	
13	v_1	5.84×10^4	2.05×10^3	235.8	48.9	15.2	6.3
	v_H	585.1	89.0	21.7	7.5	3.4	2
	v_2	806*	85.1*	19.3*	7.2*	4.0*	2.9
	v_c	—	—	14	3.6	2.44	2
	N_B	512	32	8	2	2	2
14	v_1	2.05×10^5	5.56×10^3	537.6	97.6	26.8	10.0
	v_H	1.09×10^3	154.6	34.9	11.1	4.7	2.5
	v_2	1913*	160*	30.1*	9.8*	4.9*	3.3
	v_c	—	—	—	4.5	2.76	2.16
	N_B	1024	48	16	4	2	2
15	v_1	7.44×10^5	1.55×10^4	537.6	97.6	26.8	10.0
	v_H	2.05×10^3	270.8	56.9	16.9	6.6	3.3
	v_2	4656*	309.8*	48.9*	13.8*	6.2*	3.9
	v_c	—	—	—	6	3.14	2.36
	N_B	2048	—	32	4	2	2
16	v_1	2.78×10^6	4.47×10^4	3.11×10^3	431.0	94.0	28.5
	v_H	3.86×10^3	478.4	94.0	26.0	9.5	4.4
	v_2	1.16×10^4*	617.3*	82.0*	24.2	8.8	4.7
	v_c	—	—	—	9	3.67	2.6
	N_B	2048	—	32	—	2	2
17	v_1	1.07×10^7	1.33×10^5	7.81×10^3	951.5	185.2	50.7
	v_H	7.32×10^3	851.1	157.2	40.8	13.9	6.0
	v_2	2.96×10^4*	1264*	140.0*	37.5	12.1	5.9
	v_c	—	—	—	18	4.4	2.9
	N_B	4096	—	64	—	4	2

v_1 = the upper bound obtained from the "hard sphere" model.
v_2 = the improved upper bound, using the "soft sphere" model.
v_H = the Hamming bound.
v_c = the Maximal Codistant bound.
N_B = the best value actually found. These entries have been taken largely from Laemmel, except for entries marked with a double asterisk.

The maximal codistant model assumes a given n space and a number of message points, v, to be disposed in such a manner as to yield a maximal mutual separation. By maximizing the minimum of such distances, a corresponding maximum error-correcting capability will be indicated. The reasoning proceeds on the basis that, in a given n space, the sum of all combinations of all distances between the v message points (for a binary field) can not exceed $v^2n/4$. This magnitude is obtained when the ith digit of all v words is examined. To assure a maximum sum of distances (in the ith column), there must be an equal number of ones and zeros. (Hence, a necessary but insufficient condition for meeting the bound is that v be even.) This produces the sum $(v/2)^2$.

This is then increased by the factor n to account for all dimensions. The next step is to let such total distance be distributed, so that all mutual distances are equal. This assures that the minimum distance will be maximized. Since there are $\binom{v}{2}$ combinations of "paths" among v points, the minimum distance between message points cannot exceed $nv/[2(v-1)]$. Hence, $e_c \leq nv/[4(v-1)] - \frac{1}{2}$ would be the upper limit for error correction. It will be observed that the error-detection upper bound, $e_d \leq nv/[4(v-1)]$, can be realized exactly in many existing codes. v need not necessarily be a power of 2 (as in a systematic code) for meeting the bound. (An example is $n = 11$, $v = 12$, and $e = 3$. This is achieved by using all eleven cyclic shifts of the code word, 11101101000, and the zero vector.) Note also that for reasonably large v, the well-known result of $\lim_{v\to\infty} e = n/4$ immediately becomes apparent.

As stated previously, the maximal codistant bound applies only to codes in which e is relatively large. This is plausible from the standpoint that, if the space is sufficiently large, a few message points can always be disposed to meet, almost or exactly, the codistance symmetry requirement.

The range over which the maximal bound appears to be equal or superior to

* Received by the PGIT, May 20, 1960.

[1] N. Wax, "On upper bounds for error-detecting and error-correcting codes of finite length," IRE Trans. on Information Theory, vol. IT-5, p. 168; December, 1959.

[2] R. W. Hamming, "Error-detecting and error-correcting codes," *Bell Sys. Tech. J.*, p. 147; April, 1950.

the Hamming bound is $2e + 1 \leq n \leq 4e + 1, e \neq 1$ and the trivial case, $2e + 1 \leq n \leq 4e, e = 1$.

The table in Prof. Wax's paper is herein duplicated and supplemented to include the maximal codistant bound. In this extended form, the inequality becomes $v_c \leq (4e + 2)/(4e + 2 - n)$. Again, the limits on $n = f(e)$ have been observed, which limits the entries only to part of the table.

In order to highlight the ranges over which each of the bounds excel, the best limit on v_i for given n and e is shown in Table II.

TABLE II

		e					
		1	2	3	4	5	6
n	7	H	S	HC			
	8	S	C	C			
	9	S	S	C	HC		
	10	S	S	C	C		
	11	S	S	C	C	HC	
	12	H	S	C	C	C	
	13	H	S	C	C	C	HC
	14	H	H	S	C	C	C
	15	H	H	S	C	C	C
	16	H	H	S	C	C	C
	17	H	H	S	C	C	C

H—Hamming's Bound.
S—Wax's Soft Sphere Bound.
C—Maximal Codistant Bound.
HC—indicates that both bounds yield the same result.

JAMES P. LIPP
General Electric Co.
Advanced Electronics Ctr.
Ithaca, N. Y.

A Note on Error Statistics in Fading Radioteletype Circuits*

A great deal of study has been devoted in recent years to the problem of detecting binary data which are transmitted through noisy, fading channels. Nearly all published theoretical investigations, however, have been concerned only with minimal first-order error probabilities. There is at present considerable interest in applying error-correcting codes to fading radio circuits, and this raises the question of higher-order error probabilities. A sample computation of one such statistic is given below, and its relevance is discussed briefly.

Reiger and others have shown[1] that when additive white Gaussian noise is the only form of disturbance, the error probability in noncoherent, matched-filter detection of frequency-shift-keyed teletype is

$$P(\epsilon_1 \mid r_1) = \tfrac{1}{2} \exp\left[-\frac{r_1}{2}\right] \qquad (1)$$

where r_1 is the signal-to-noise power ratio.

* Received by the PGIT, February 15, 1960. The work discussed in this letter was done in 1957–1958 at the Long Range Radio Branch of the USASRDL, Fort Monmouth, N. J.

[1] S. Reiger, "Error probabilities of binary data transmission systems in the presence of random noise," 1953 IRE NATIONAL CONVENTION RECORD, pt. 8, pp. 72–79.

If the signal is subject to nonselective fading, r_1 must be considered as a random variable and a mean value for the error probability can be found by averaging over all values of r_1. For fading which is Rayleigh-distributed, the probability density function (pdf) of r_1 is

$$p(r_1) = \frac{1}{\rho} \exp\left[-\frac{r_1}{\rho}\right], 0 \leq r_1 \qquad (2)$$

where ρ is the mean value of the fading signal/noise ratio. Thus the mean error probability under Rayleigh-fading conditions is[2]

$$P(\epsilon_1) = \int_0^\infty p(r_1) P(\epsilon_1 \mid r_1)\, dr_1 = \frac{1}{\rho + 2} \qquad (3)$$

Clearly, this result applies only to sets of decisions whose members are spaced sufficiently far apart in time, frequency, or space so that they are essentially independent; it certainly does not apply to contiguous data pulses which are transmitted at a rate considerably faster than the "speed" of the fading.

To evaluate the average first-order conditional error probability in slow-Rayleigh-fading situations, one must find the average probability of two adjacent errors. This probability, given the signal-to-noise ratios at the two detection instants, is

$$P(\epsilon_1, \epsilon_2 \mid r_1, r_2) = \tfrac{1}{2} \exp\left[-\frac{r_1}{2}\right] \cdot \tfrac{1}{2} \exp\left[-\frac{r_2}{2}\right]. \qquad (4)$$

An average value for this error-doublet probability can be found by weighting (4) with the bivariate pdf of r_1 and r_2, which is[3]

$$p(r_1, r_2,) = \frac{1}{\rho^2(1 - \varphi^2)} \cdot \exp\left[-\frac{r_1 + r_2}{\rho(1 - \varphi^2)}\right] \cdot I_0\left[\left(\frac{2\varphi}{1 - \varphi^2}\right)\sqrt{\frac{r_1 r_2}{\rho^2}}\right] \quad \begin{cases} 0 \leq r_1, r_2 \\ 0 \leq \varphi \leq 1 \end{cases}. \qquad (5)$$

(The parameter φ which appears in (5) is the correlation coefficient over the pulse length of the quadrature components of the fading signal. In essence, φ is a measure of the "speed" of the fading. When $\phi \approx 0$, the fading rate exceeds the data rate, while when $\varphi \approx 1$, the fading is very slow indeed.) The average probability of two adjacent errors is

$$P(\epsilon_1, \epsilon_2) = \int_0^\infty \int_0^\infty p(r_1, r_2) \cdot P(\epsilon_1, \epsilon_2 \mid r_1, r_2)\, dr_1\, dr_2 = \frac{1}{\rho^2(1 - \varphi^2) + 4\rho + 4}. \qquad (6$$

[2] M. Masonson, "Binary transmissions through noise and fading," 1957 IRE NATIONAL CONVENTION RECORD, pt. 2, pp. 69–83.

[3] S. O. Rice, "Mathematical analysis of random noise," in "Selected Papers on Noise and Stochastic Processes," N. Wax, ed., Dover Publications, Inc., New York, N. Y., pp. 133–294; 1954.

The average first-order conditional erro probability is

$$P(\epsilon_2 \mid \epsilon_1) = \frac{\rho + 2}{\rho^2(1 - \varphi^2) + 4\rho + 4}. \qquad (7$$

When $\varphi \approx 0$ (fast-fading), the conditiona error probability approaches the first-orde error probability given by (3). This is condition which violates the assumption underlying (1), and which is almost neve met in practice. When $\varphi > 0$, the con ditional probability always exceeds th first-order probability. It is interesting t put some real-life numbers into (3) an (7). For $\rho = 30$ db and $\varphi = 0.995$, whic are values characteristic of high-frequenc long-range radio practice:

$$P(\epsilon_1) \simeq 10^{-3}$$

$$P(\epsilon_2 \mid \epsilon_1) \simeq 7 \times 10^{-2}.$$

These values imply that error-burst-cor recting codes may be useful for fadin radio circuits.

The preceding results are the simples of many which can be derived. Recurrenc formulas can sometimes be develope (depending on the correlation character istics of the fading) to reduce the multipl integrals which arise in evaluating multiple error clusters, and combinatorial analysi can be used to determine what types o error-clusters warrant consideration withi given redundancy constraints. There are however, serious limitations on the result given here, and on direct extensions thereto For example, (6) and (7) are valid onl for white noise interference and Rayleig fading, and apply only to nondiversit reception. Furthermore, they are limite to FSK systems with narrow frequenc shifts, for an underlying assumption i that fades in the "mark" and "space sub-channels are completely correlated. I this assumption is not made, tractabl general results can't be derived readil because the conditional probabilities wil depend upon what character is being sent Unfortunately, the two classes of servic wherein FSK is most widely used—h layer propagation and ionoscatter propaga tion—do not conform to these limitation In the former, stationary white noise i an inadequate model for additive dis turbances, and in the latter, wide frequenc shifts are normally used. Thus, calcula tions of the sort discussed here may prov to be mainly of academic interest.

The writer wishes to thank officials a the USASRDL and the Department o Defense for permission to publish thi letter.

H. VOELCKE
The Imperial College of Sc
and Tech
London, Eng

Contributors

Norman Abramson (S'55—M'59) was born on April 1, 1932, in Boston, Mass. He received the B.A. degree in physics from Harvard University, Cambridge, Mass., in 1953, the M.A. degree in physics from the University of California at Los Angeles, in 1955, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, Calif., in 1958.

He was employed part-time from 1953 to 1956 as a research engineer in the systems analysis laboratory of Hughes Aircraft Company, Los Angeles, Calif. In 1955 he entered the electrical engineering department at Stanford under the Consolidated Engineering Fellowship; he is now an assistant professor of electrical engineering at Stanford. He has worked in the analysis of missile guidance systems, binary coding theory, binary communications systems and the applications of statistical decision theory.

Prof. Abramson is a member of Sigma Xi.

❖

Carmen N. Campopiano (M'56) was born in Trenton, N. J., on February 1, 1924. He received the B.S. and the M.S. degrees in mathematics in 1949 and 1951, respectively, from Rutgers University, New Brunswick, N. J.

From 1949 to 1952, he was a teaching assistant in the Mathematics Department of Rutgers University. From 1952 to 1954, he was engaged in operations research and weapon control systems at Sperry Gyroscope Company, Great Neck, L. I., N. Y. In 1955, he was a member of the mathematics section of the Nuclear Power Division of Combustion Engineering Inc., New York, N. Y. From 1956 to 1958, he was a senior research associate with Microwave Research Institute, Polytechnic Institute of Brooklyn, Brooklyn, N. Y., where he investigated coding for facsimile transmission and the diffraction of electromagnetic waves. He is presently working in the field of digital communication in the Airborne Systems Division of Radio Corporation of America, Camden, N. J.

Mr. Campopiano is a member of the American Mathematical Society, Mathematical Association of America, Society for Industrial and Applied Mathematics, and Sigma Xi.

❖

Carl W. Helstrom was born in Easton, Pa., on February 22, 1925. From 1944 to 1946 he served as a radio technician in the United States Navy. In 1947 he received the B.S. degree in engineering physics from Lehigh University, Bethlehem, Pa., and in 1951 the Ph.D. degree in physics from the California Institute of Technology, Pasadena.

Since 1951 he has been employed at the Westinghouse Research Laboratories, Pittsburgh, Pa., where he is now an advisory mathematician.

Dr. Helstrom is a member of Phi Beta Kappa and the American Physical Society.

❖

Morris J. Levin (S'50—A'53—M'57) was born in Burlington, Vermont, on June 21, 1928. He received the B.S. degree from the University of Vermont, Burlington, in 1950, the M.S. degree from M.I.T., Cambridge, Mass., in 1952, and the Doctor of Engineering Science degree from Columbia University, New York, N. Y., in 1959, all in electrical engineering.

He has been employed by Lincoln Laboratories, Lexington, Mass., and the Atomic Instrument Company, Cambridge, Mass., and has served in the United States Army. At present he is a member of the Applied Research Group of the RCA Missile Electronics and Controls Division, Burlington, Mass.

Dr. Levin is a member of Sigma Xi, Phi Beta Kappa, the Society for Industrial and Applied Mathematics, and the Institute of Mathematical Statistics.

Richard A. Silverman (M'54—SM'58), for a photograph and biography please see page 507 of the September, 1960, issue of these TRANSACTIONS.

❖

Charles A. Stutt (S'43—A'46—M'55) was born in Avoca, Neb., on November 12, 1921. He received the B.S. degree from the University of Nebraska, Lincoln, in 1944, and the D.Sc. degree from the Massachusetts Institute of Technology, Cambridge, in 1951, both in electrical engineering.

From 1944 to 1946, he was employed in the research department of the Stromberg-Carlson Company, Rochester, N. Y. In 1947 and 1948, he was a recipient of an industrial electronics fellowship at M.I.T.; and from 1948 to 1951 he served there as research assistant in the Research Laboratory of Electronics and instructor in the Electrical Engineering Department. In 1951, he was in England participating in transatlantic FM studies conducted by the Research Laboratory of Electronics. In 1952 he joined the M.I.T. Lincoln Laboratory, where he was engaged principally in radar and long-range communication research. In 1957, he joined the information studies section of the General Electric Research Laboratory, Schenectady, N. Y., where he is presently engaged in communications research.

Dr. Stutt is a member of Sigma Tau, Pi Mu Epsilon, and Sigma Xi.

❖

W. M. Wonham was born in Montreal, Can., on November 1, 1934. He received the B.E. degree in engineering physics from McGill University in 1956.

Since September, 1956, he has been working toward the doctoral degree in the department of control engineering at the University of Cambridge, Cambridge, England. His research is concerned with the analysis and optimization of relay and saturating control systems with random inputs.

Abstracts

This section of the issue is devoted to abstracts of material which may be of interest to PGIT members. Sources are Government, Industrial and University reports, and books and journals published outside the United States. Readers familiar with material of this nature which is suitable for abstracting are requested to communicate the pertinent information to one of the Editors or Correspondents listed below.

Chain Codes and their Electronic Applications—F. G. Heath and M. W. Gribble (in English). (IEE Monograph No. 392M; July, 1960.)

A "chain code" is one using fixed-length groups of n binary digits, where each group uses $n-1$ consecutive digits of the previous group shifted along one place, and one new digit constructed by some logical rule from the digits of the previous group. The groups can be constructed consecutively by a combination of a shift register with logic circuits for constructing the one new digit from the old ones. This method of construction bears some resemblance to the construction of check digits in Hamming codes, and chain codes can be used for error detection and for error correction on the basis of exhaustive comparison and minimum number of erroneous digits.

Chain-code generators also serve as counters, or for frequency division of variable-frequency input pulses.

Distortion Distribution and Error Probability in Binary Transmission Systems—Y. Hoshiko, T. Minami, and T. Omori (in Japanese). (*J. Inst. Elec. Comm. Engrs. Japan*, vol. 43, pp. 146–153; February, 1960.)

The most important factors which cause error in binary transmission systems are noise and transmission distortion. In this paper, we discuss the residual response (intersymbol interference) at sampling points and analyze the general form of the distribution function of the residual response of a random binary signal pattern. This distribution function is ultimately given in terms of the characteristic transmission distortion function. The over-all effect of both transmission distortion and noise can then be obtained by convolution of distribution functions. A numerical example is given using the distortion of a constant-K filter circuit, and the error rate is computed at various signal-to-noise ratios.

On $(\sin x)/x$ Sampling in the Case of Non-Band-Limited Functions—K. R. Johnson (in English). (Lincoln Lab., Mass. Inst. Tech., Lexington, Mass., Tech. Rept. 195; February 16, 1959.)

A theoretical investigation is presented of the use of $(\sin x)/x$ sampling to represent a function $f(t)$ having a Fourier transform $g(\omega)$ such that for large $|\omega|$, $|g|$ and $|dg/d\omega|$ tend to zero as $|\omega|^{-\alpha_1}$ and $|\omega|^{-\alpha_2}$, respectively, when $\alpha_1 > 2$, $\alpha_2 > 2$. An upper bound is given for the error

$$\epsilon = \left| f(t) - \sum_{h=-\infty}^{\infty} f\left(\frac{\pi n}{\omega_s}\right)\frac{\sin(\omega_s t - \pi n)}{\omega_s t - \pi n} \right|$$

as a function of ω_s and parameters describing the behavior of the transform of f. It is shown that under the above indicated conditions ϵ approaches zero uniformly in t as ω_s approaches infinity. Also, an upper bound is obtained for ϵ for the case of summation over only a finite number of sample points. The uniformity of the convergence of the representation as the number of sample points increases is investigated. In addition, there is a rigorous proof of the $(\sin x)/x$ sampling theorem for band-limiting continuous function in L_1.

Representations of Vector-Valued Random Processes—E. J. Kelly and W. L. Root (in English). (Lincoln Lab., Mass. Inst. Tech., Lexington, Mass., Group Rept. 55-21; March 7, 1960.)

In this note we extend two representation theorems, well-known for ordinary random processes, to the case of vector-valued random processes which are continuous in mean square. The process to be represented may have finitely or, if certain convergence conditions are satisfied, infinitely many components. The first of these extensions is an orthogonal series representation which holds on a finite interval. It is closely analogous to the Karhunen-Loève representation for complex-valued random processes, and we prove its validity by an

rgument which parallels a standard proof f the Karhunen-Loève theorem. A eneralization of Mercer's theorem for ontinuous, definite kernels is required, ut it also can be proved by the same rguments as in the classical case.

The second generalization is a vector pectral representation theorem, holding n the entire real line, under the weakest atural assumption, namely, harmonizbility. This is not simply a reapplication f the usual theorem to a different space, ut, like the orthogonal series representaion, it requires a slight extension of the revious results.

The representations derived here promise o be useful in signal-detection-in-noise roblems where the detecting apparatus as more than one sensor, as for example he analysis of a "space-diversity" radio r radar receiver with multiple antennas.

Error Probabilities for the Ideal Detecion of Signals Perturbed by Scatter and Noise—R. Price (in English). (Lincoln Lab., Mass. Inst. Tech., Lexington, Mass., Group Rept. 34–40; October 3, 1955.)

Previous studies have shown how the unctional form of the ideal, probabilityomputing receiver may be found for a ommunication system in which one of a umber of possible bandpass waveforms is ransmitted over a channel containing one r more scatter-paths and additive white Gaussian noise. In this paper, a measure of he over-all performance of such a system, mploying the ideal receiver, is found by ssuming that a decision about which waveform was transmitted is made on the asis of the *a posteriori* probabilities computed by the receiver. The probability of rror $P_m(e)$ in this decision is taken as the system performance, and, since it is the ninimum probability of error obtainable y any receiver for the given transmitter and channel, this performance measure may be considered to be a property of the ransmitter-channel combination alone.

Only for the rather restricted case of n-off transmission through a single path ave solutions been obtained for $P_m(e)$, and even then only in terms of the eigenvalues of a homogeneous integral equation nvolving the envelope of the transmitted waveform and a function that yields the complete statistical description of the scatter-path. In cases where the path is of a very simple Markoff type and the envelope is either constant for a finite interval or exponentially decaying over an infinite time interval, the integral equation has been solved, and specific numerical results for $P_m(e)$ have been obtained after rather laborious computation. In the limiting cases of very slow and very fast scatter-path fluctuations, it has been possible to obtain good approximations to $P_m(e)$ for a wider variety of envelopes and scatter-path statistics.

In the special cases previously mentioned, sufficient results have been obtained to demonstrate two interesting characteristics of scatter-path transmission which may hold true in general. First, for sufficiently small additive noise, fast scatter-path fluctuations yield better system performance than slower ones. Second, in the case of a constant transmitted envelope, increasing the duration of the transmitted waveform in the same ratio that the additive noise is decreased leaves $P_m(e)$ unchanged for slower scatter fluctuations, while the duration varies as the square of the noise for fixed performance when the scatter is fluctuating rapidly.

It has been possible to draw some comparisons between the performance of the system employing the ideal receiver and the best performance obtained from systems having identical transmitter-channel combinations but different receivers. In one case, the performance has been found for a rather fictitious receiver that is somehow supplied directly with complete information about the scatter-path fluctuations as well as with the received waveform. Naturally, this performance is better than that obtained with the more realizable ideal receiver, and this improvement is shown to increase with faster fluctuation of the scatter-path. Two other receivers that have been studied employ a correlation-envelope detector and an energy-measurement detector, respectively. The former is shown to approach ideal performance for slow scatter fluctuations, while the latter approaches the ideal for fast scatter fluctuations.

An incidental result of this analysis is the determination of the probability distributions of finite-time energy measurements made on narrow-band Gaussian noises. These distributions are given in an appendix, together with details of thier computation.

A large portion of the work carried out in this paper applies directly to the interesting problem of determining with greatest efficiency whether or not a noise-like, Gaussian signal, or narrow bandwidth is present in a background of additive white Gaussian noise. Such problems are encountered in radio-astronomy and molecular spectroscopy.

Some Aspects of the Relative Efficiencies of Indian Languages—B. S. Ramakrishna, *et al.* (in English). (Dept. of Elec. Comm. Engrs. Ind. Inst. Sci., Bangalore 12, South India, Tech. Rept. 1; 1959–1960.)

This monograph describes some recent statistical studies in the Indian languages Hindi, Marathi, Tamil, Malayalam, Telugu, and Kannada. The keynote of this study is that where different languages can be employed to serve the same end purpose, their relative performances as alternate means of communication can be compared on the basis of the principles of information theory. Although this report is restricted mainly to the Indian languages, the concepts and techniques developed may find wider application.

The study is based on two notions: The first of these is that for messages in different languages having the same meaning, their total entropies form inverse measures of the relative efficiencies with which the languages encode semantic content into linguistic symbols. Using samples of texts in one language and their translations in another as semantically-equivalent materials for comparison, the authors develop an information theoretical model of translation as a process which leaves the semantic content (but not the entropy) invariant, and carries with it at the same time a certain noise in the semantic sense. On the basis of this model, the number of selective bits of information in several Indian languages semantically equivalent to one bit in English have been obtained. The implication of these results in respect of the choice of a suitable language for telegraphic communication is also discussed.

The second notion introduced here considers different scripts as alternate means of transcribing phonetic content into written symbols. It is argued (with examples chosen from the history of writing) that the speed with which a script can be written is one of the significant criteria of its merit. In the terminology of information theory this amounts to saying that between two scripts the one which takes a lesser time to write a given phonetic content has a larger channel capacity. The question whether the adaption of Roman script to Indian languages in place of their current scripts based on syllabary results in speedier or slower writing of these languages is then considered.

A subjective test was designed as follows: A number of individuals who could write both the Roman and an Indian script were asked to copy a given set of nonsense syllables in both the scripts, and the times taken by them were measured. An analysis of variance was then performed to arrive at an estimate of the intrinsic relative time requirements of the Indian and Roman scripts independently of the practice of the individuals in the two scripts. The indications of the preliminary results are that Tamil, Kannada, and Hindi may save about 5 per cent of the time by using Roman script, while Telugu and Malayalam may take about 5 per cent more time when written in Roman script.

The monograph also gives information concerning the relative frequencies of the different speech sounds, different types of syllables, and the methods of computation employed.

Codes with Zero Correlation—D. N. Tompkins (in English). (Engrg. Division, Hughes Aircraft Co., Culver City, Calif., Tech. Memo. 651; June, 1960.)

Codes of ternary digits are presented in which each code possesses zero cyclic autocorrelation except when perfectly correlated. The lengths of these codes are up to 19 digits.

The codes are found by a technique which generates lists of binary sequences. Each list is prime in the sense that no sequence in the list is a cyclic permutation, an order inverse or an amplitude inverse of any other sequence in the list. It is shown that the size of such a list approaches an exponential behavior as the lengths of the sequences are increased.

Application of the codes to specific areas such as pulse compression, Doppler meas-

urements, and noise synthesis is discussed. Since cyclic rather than linear correlation is requisite, its implementation is also presented.

On the Detection and Estimation Problem for Multiple Nonstationary Random Processes—J. K. Wolf (in English). (Dept. of Elec. Engrg., Princeton University, Princeton, N. J., Ph.D. dissertation; October, 1959.)

Two major problems in the study of statistical communication theory are the detection and estimation of signals corrupted by additive noise. In this investigation, certain aspects of both problems are examined with particular emphasis on the processing of multiple, nonstationary time series, *i.e.*, correlated random processes, the statistics of which vary with time.

As a preliminary to this study, some properties of random processes are examined, including an introduction to the concepts of stationarity, uniformity, ergodicity, and time-varying power spectra for nonstationary processes. Some theorems related to the optimum filtering of nonstationary signals are also examined.

A general formulation of the multidimensional linear filtering problem is presented next and is used in finding the minimum mean-squared error, linear time-varying filters for multidimensional inputs. Both polynomial signals and signals which are sample functions of random time series are considered. The solution to certain types of matrix integral equations which occur in this work are discussed.

The subject of random systems is examined from two different viewpoints. In the first case, a study is made of the filtering of signals which have been transmitted through a linear filter with certain random transmission characteristics. Both time-varying linear filters with random parameters and linear filters with impulsive responses which are sample functions of random processes are discussed. In the second case, the design of optimum systems is studied when the components used to construct these systems have values which vary from their nominal values in some random fashion. It is shown that the optimum nominal values of such non-ideal components will differ from the values of perfect components except under certain special conditions.

Next, the problem of detecting signals in noise is considered for the multiple input model, where each of the inputs can contain one out of many possible signals. The detection procedure for this model becomes, in general, the testing of multiple hypotheses. Two detection criteria are examined for choosing between multiple hypotheses and it is found that for both criteria, the decision is based upon the calculation of the likelihood functions for the various signals. Systems for calculating these likelihood ratios are then examined for deterministic signals, with and without random parameters, and for signals which are sample functions of random processes. A multidimensional matched filter is introduced, and its relationship to the detection problem is examined. The choice of signals which minimizes the probability of a wrong decision is found for the case when there can be only two possible signal combinations at the input.

Finally, a useful expansion is presented for stochastic processes which are functions of two variables. The utility of this expansion in specifying optical filters for the detection of signals in noise is examined. The defining relationship for an optical matched filter is derived and related to this optical detection problem.

Vector Stochastic Processes in Problems of Communication Theory—E. Wong (in English). (Dept. of Elec. Engrg., Princeton University, Princeton, N. J., Ph.D dissertation; May, 1959.)

Some aspects of the application of multidimensional stochastic processes to communication theory are studied. A standard notation and a survey of mathematical techniques are introduced at the outset, in order that continuity and consistency be maintained in the remainder of the text.

The prediction and filtering of multiple stationary time series are studied. Particular emphasis is given to the methods of solution. The matrix factorization procedure due to Wiener and Masani is extended to include the continuous case. An intuitive derivation analogous to that of Bode and Shannon is given. A class of two-dimensional problems is defined where the factorization of the spectral density may be avoided.

The maximum likelihood estimation of continuously modulated vector Gaussian processes is derived. The derivation uses a multidimensional orthogonal expansion due to others. Two practical schemes of modulation, quadrature modulation and single sideband, are discussed in this framework.

The joint probability distribution of quadratic functionals involving vector Gaussian processes is studied. The characteristic function for the joint distribution is found in terms of the eigenvalues of a homogeneous matrix integral equation. In special cases, this problem has been studied by others. Reduction to their results is demonstrated.

A relationship between the Fokker-Planck equation and an expansion of second-order probability density functions is developed. This relationship is shown to define a class of stationary Markoff processes which have useful properties.

Two problems are considered where disturbances in addition to additive noise play an important role in the reliability of communication. First, the problem where the signal has been passed through a network with random parameters is analyzed. Second, the effect of uncertainties in connections is considered. In both cases the minimum mean-squared error filters are derived.

The following papers were publishe singly by the Professional Group on In formation Theory (I) and the Professiona Group on Automata and Automatic Contro (A) of the Institute of Electrical Communi cation Engineers of Japan, 2–8, Fujimicho Chiyodaku, Tokyo, Japan. All are i Japanese, but English abstracts are give below when available. The affiliation o the author is given so that intereste readers may contact the author directl for further information.

On a Distance-Preserving Code Syster (I; May 20, 1960)—M. Fujii. (Electro technical Lab., 1, 2-chome, Nagata-cho Chiyuda-ku, Tokyo.)

This code system is constructed of finite number (n) of codes, expressed b binary symbols, and is characterized a follows: the codes are ordered in th manner of n successive integers; and for given positive integer not greater than n, 1) if the value of the ordinary distanc between two integers (mentioned above is smaller than the given integer, the this distance is equal to Hamming distanc between codes corresponding to both integers, and 2) if the former is not smalle than the given integer, then it can only b seen that the latter is also not smalle than the given integer.

Such a code system as mentioned above is devised for a treatment of the information in pattern recognition. The conception o the restricted-distance preserving code system is explained and a method of its construction is clarified.

Zero-Crossing Information of an I Speech Wave and its Band-Compression Systems (I; June 24, 1960)—K. Hirmatzu. (Tokyo Electrotechnical College, 2–2 Kanda Nishiki-cho, Chiyoda-ku, Tokyo.)

The syllable articulation of "carrier leaked" SSB clipping is as high as that o SSB, approximately 90 per cent. Hence carrier-leaked SSB clipping is available as a new information source for communi cation systems.

This paper indicates that the zero crossing information in carrier-leaked SSB is a kind of special sampling of the speech wave, and that the carrier-leak level give the properties of this information. Namely the information in low-level carrier-leaked SSB clipping is similar to that of narrow band FM, but that of high-level, to wide band FM. One is applicable to a speech level dynamic control system, the other t various speech-parameter extraction o band-compression systems.

Optimum Recognition Systems for Mark off Processes (A; July 7, 1960)—K. Horiu chi. (Waseda University, School of Scienc and Engrg., 1–647, Totsuka-cho, Shinjuku ku, Tokyo.)

Every automaton has a system or orga of perception in a broad sense and recognize

e perceived objects by its allowable nction. In practice, the individual percept dergoes some distortions or deteriora- ns from the conceptional pattern stored advance in the brain or in some equiva- nt system for recognition. Furthermore, en if all the given percepts are exactly e same as the conceptional ones, the evitable imperfection of the organ gives e to some deteriorations of the perceived tterns. In this connection, there exists ambiguity of recognition, and we can roduce probability theory into this field.

For most cases, the recognized objects given as the time series of a Markoff ocess. Therefore, any optimum system ich recognizes such objects with mini- m error can be constructed by means suitable considerations of the prob- ility properties of these time series.

In this report a general treatment of cognition, a system in which the expected k of misrecognition is minimum for a of losses preassigned to each allowable se, and a system with minimum prob- ility of misrecognition are studied in tail from the standpoint of statistical cision theory for the Markoff process.

The optimum systems consist of select- r switch circuits with some memory, nsitional probability calculators, ampli- rs, comparison circuits, etc. Explicit amples are illustrated for electric circuits.

Error Rates in Binary Pulse Transmis- n (I; April 22, 1960)—Y. Hoshiko and Sugiyama. (Electrical Communication b., 1551 Kichijoji, Musashino-shi, kyo.)

The error in a binary transmission system caused chiefly by the transmission tortion and the noise. As for the former, probability-theoretical problem of the ersymbol interference is a very compli- ed one, and hitherto no analysis has n made. Here we treat the analysis this problem, then consider the effect the noise, and finally calculate the over- error probability in the system.

The Normalization of Patterns (I; July 15, 1960)—S. Inomata. (The Electro-technical Lab., 1, 2-chome, Nagata-cho, Chiyuda-ku, Tokyo.)

A computational algorithm which normalizes a two-dimensional real pattern with respect to its intensity, position, size and rotation has been proposed.

After a "center of gravity" of the pattern, analogous to that in mechanics, is introduced, a two-stage coordinate transformation, one Cartesian, another polar, is performed, yielding a favorable representation of the pattern for the normalization. Then a finite Fourier transform of the pattern represented by polar coordinates is taken with respect to the angular variable, giving rise to a function-series representation of the pattern. Next, a logarithmic coordinate conversion is introduced as to the radial variable, and an infinite Fourier transform in the complex domain is taken, giving also a function-series representation.

The final normalized expression of the pattern is obtained by a simple algorithm composed of an infinite integral and a division involving the above function series which vanishes rather rapidly for "smooth" patterns for both large series numbers and large arguments, and reduces to "one" function in case the pattern is point symmetrical.

Lastly, the computational flow-chart for the normalization is described and the necessary machine time is estimated, although approximately.

Musical Composition and Performance by Means of Computing Machines (A; April 21, 1960)—H. Isida. (Faculty of Science, Tokyo University, Bunkyo-ku, Tokyo.)

Synchronization for Multi-channel PCM (I; April 22, 1960)—Y. Nakamaru and H. Kaneko. (Nippon Electric Co., 1753 Tamagawamuko, Shimonuma-gun, Kawasaki-shi.)

The digital method of synchronization is suitable for framing of digital transmission channels such as long-frame PCM. In this paper a new method of synchronization with an excellent characteristic is derived from a theoretical analysis of the recovery process. In this system the framing-pulse sequence is located in a group at the beginning of a frame, and a hunting procedure is performed in such a way as to reset immediately the channel separator at each instant of error detection.

It is shown that according to this method, normal synchronism can usually be restored in only one frame of collapse by employing ten or more framing pulses in a frame. Optimum codes and a method of frame composition which provides an optimum characteristic are presented. The stability of synchronism is highly improved by adding a locking circuit. The results of the analysis are proved experimentally by using an empirically made recovery process simulator.

On the Sequential Design of Experiments (A; July 7, 1960—M. Sakaguchi. (University of Electro-communications, 14 Kojima-cho, Chofu-shi, Tokyo-to.)

On the Application of a Speech Clipper to an SSB Radio Telephone System using VODAS (I; July 15, 1960)—U. Tsuruoka and J. Nakamichi. (Japan's Overseas Radio and Cable System, 1–5 Otemachi, Chiyoda-ku, Tokyo.)

This paper deals with signal-to-noise improvements by the use of a speech clipper in an SSB radio telephone system using the VODAS. The results of the studies show: 1) it is possible to clip the speech currents 10 db from the peak level at the input of the transmitter with little influence upon the naturalness of the voice; 2) in the application of a speech clipper to a system using the VODAS, the intelligibility of the system can be improved to be more than 93 per cent; this figure is that observed in one-way transmission, and equivalently about ten times the power of the usual transmitter can be obtained.

Index to

IRE TRANSACTIONS

ON

INFORMATION THEORY

Vol. IT-6, 1960

IRE Transactions on Information Theory

Index to Volume IT-6, 1960

CONTENTS

Volume IT-6, Number 1, March, 1960

Volume IT-6, Number 2, Special Monograph Issue, April, 1960

Volume IT-6, Number 3, June, 1960

Volume IT-6, Number 4, September, 1960

Volume IT-6, Number 5, December, 1960

Index to Authors

(Listed by month and page.)

Index to Subjects

(Listed by month and page.)

Index to Footnote References

The following is an alphabetical compilation of all footnote and bibliographical references published in PGIT TRANSACTIONS *papers and correspondence during the year. The month and page where each reference was listed are given in parentheses.*

A

Abramson, N., "The application of 'comparison of experiments' to detection problems," 1958 IRE NATIONAL CONVENTION RECORD, pt. 4, pp. 22–26. (Dec 530)
Abramson, N., "Application of 'Comparison of Experiments' to Radar Detection and Coding Problems," Electronics Labs., Stanford University, Stanford, Calif., Tech. Rept. No. 41; July 28, 1958. (Dec 530)
Abramson, N., "A class of systematic codes for non-independent-errors," IRE TRANS. ON INFORMATION THEORY, vol. IT-5, pp. 150–157; December, 1959. (Mar 13, 16) (Sep 502)
Abramson, N., "A Class of Systematic Codes for Non-Independent Errors," Electronics Labs., Stanford University, Stanford, Calif., Rept. No. 51; December 30, 1958. (Mar 13)
Abramson, N., and Elspas, B., "Double-Error-Correcting Encoders and Decoders for Non-Independent Binary Errors," presented at Internatl. Conf. on Information Processing, Paris, France; June 15–22, 1959. (Mar 13)
Aitken, A. C., "On least square and linear combination of prediction," *Proc. Roy. Soc.* (*Edinburgh*), vol. 55, pp. 42–48; November, 1934. (Sep 484)
Agathen, B.: See Sutro, P. J.
Alberth, A. A., "Fundamental Concepts of Higher Algeb
University of Chicago Press, Chicago, Ill.; 1956. (Mar 18) (
451, 453)
Allis, W. P., and Herlin, M. A., 'Thermodynamics and Statist
Mechanics," McGraw-Hill Book Co., Inc., New York, N.
1952. (Sep 437)
Anderson, V. C., "Deltic Correlator," Harvard Acoustics L
Harvard University, Cambridge, Mass., Tech. Memo. No.
January, 1956. (Jun 348)
Arens, R., "Complex processes for envelopes of normal noi
IRE TRANS. ON INFORMATION THEORY, vol. IT-3, pp. 204–
September, 1957.(Jun 330)
Ashby, R. M., Josephson, V., and Sydoriak, S., "Signal Thresh
Studies," Nav. Res. Lab., Rept. No. R-3007; December 1, 1
(Apr 144, 266)

B

Baghdady, E., Ed., "Lectures on Signal Analysis," McGraw-
Book Co., Inc., New York, N. Y., ch. 10; to be published. (Jun
Balakrishnan, A. V., and Drenick, R., "On optimum nonli

extraction and coding filters," IRE Trans. on Information Theory, vol. IT-2, pp. 166–173; September, 1956. (Sep 500)
arker, R. H., "Group synchronizing of binary digital systems," in "Communication Theory," W. Jackson, Ed., Academic Press, New York, N. Y.; 1953. (Jun 328)
arrett, J. F.: See Lampard, D. G.
artlett, M. S., "An Introduction to Stochastic Processes," Cambridge University Press, Cambridge, England; 1955. (Dec 540)
asore, B. L., "Regular Simplex Coding," The Dikewood Corp., Scientific Rept. No. 1, Contract No. AF 19(604)-4078, AFCRC TN-59-165, ASTIA Doc. No. AD-213609. (Dec 516)
atchelor, G. K., "The Theory of Homogeneous Turbulence," Cambridge University Press, Cambridge, England; 1953. (Dec 548)
attin, R. H.: See Laning, J. H.
ell, D. A., "Information Theory and its Engineering Applications," Pitman Publishing Co., New York, N. Y., 2nd ed.; 1956. (Jun 410)
ello, P.: See Keilson, J.
ennett, W. R., "Response of a linear rectifier to signal and noise," *J. Acoust. Soc. Am.*, vol. 15, pp. 164–172; January, 1944. (Apr 265)
ernstein, R. I.: See Miller, K. S.
irdsall, T. G.: See Peterson, W. W.
irkoff, G., and MacLane, S., "A Survey of Modern Algebra," The Macmillan Co., New York, N. Y.; 1953. (Sep 463)
lachman, N. M., "The output signal-to-noise ratio of a power-law device," *J. Appl. Phys.*, vol. 24, pp. 783–785; June, 1953. (Mar 52)
lack, H. S., "Modulation Theory," D. Van Nostrand Co., Inc., New York, N. Y.; 1953. (Sep 422)
lake, L. V., "Interim Report on Development of Model SPS-3 (XDK) Shipborne Radar," Nav. Res. Labs., Rept. No. R-3165; August, 1947. (Apr 266)
lackwell, D., "Equivalent comparisons of experiments," *Ann. Math. Stat.*, vol. 24, pp. 265–272; 1953. (Dec 530)
lackwell, D., and Girshich, M., "Theory of Games and Statistical Decisions," John Wiley and Sons, Inc., New York, N. Y.; 1954. (Dec 530)
lanc-Lapierre, A., and Fortet, R., "Théorie des Fonctions Aléatoires," Masson, Paris, France, ch. 5; 1953. (Sep 485)
lasbalg, H., "The sequential detection of a sine-wave carrier of arbitrary duty factor in Gaussian noise," IRE Trans. on Information Theory, vol. IT-3, pp. 248–256; December, 1957. (Sep 505)
lum, M., "An extension of the minimum mean square prediction theory for sampled input signals," IRE Trans. on Information Theory, vol. IT-2, pp. 176–184; September, 1956. (Sep 484)
lum, M., "Generalization of the class of nonrandom inputs of the Zadeh-Ragazzini prediction model," IRE Trans. on Information Theory, vol. IT-2, pp. 76–81; June, 1956. (Sep 426)
lumenson, L. E.: See Miller, K. S.
ond, F. E., and Cahn, C. R., On sampling the zeros of bandwidth-limited signals," IRE Trans. on Information Theory, vol. IT-4, pp. 110–113; September, 1958. (Mar 51)
ooton, R. C., Jr., "An optimization theory for time-varying linear systems with nonstationary statistical inputs," Proc. IRE, vol. 40, pp. 977–981; August, 1952. (Sep 420)
ooton, R. C., Jr., and Goldstein, M. H., Jr., "The design and optimization of synchronous demodulators," 1957 WESCON Convention Record, pt. 2, pp. 154–170. (Sep 420)
orn, M., and Wolf. E., "Principles of Optics," Pergamon Press, Inc., New York, N. Y.; 1959. (Jun 387)
ose, A. G., "A Theory of Nonlinear Systems," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Tech. Rept. No. 309; 1956. (Sep 500)
ose, R. C., and Kuebler, R. R., Jr., "A geometry of binary sequences associated with group alphabets in information theory," *Ann. Math. Stat.*, vol. 31, pp. 113–139; March, 1960. (Dec 552)
ose, R. C., and Ray-Chaudhuri, D. K., "On a class of error-correcting binary group codes," *Information and Control,* to be published. (Sep 459)
radt, R., and Karlin, S., "On the design and comparison of certain dichotomous experiments," *Ann. Math. Stat.*, vol. 27, pp. 390–409; 1956. (Dec 530)
rennen, D. G., "On the maximum signal-to-noise ratio realizable from several noisy signals," Proc. IRE, vol. 43, p. 1530; October, 1955. (Jun 328)
rill, E. R., The Improvement in Minimum Detectable Signal in Noise Through the Use of the Long Afterglow CR Tube, and Through Photographic Integration," Radio Res. Lab., Harvard University, Cambridge, Mass., Rept. No. 411-484; February, 1944. (Apr 266)
Brillouin, L., "Science and Information Theory," Academic Press, Inc., New York, N. Y.; 1956. (Sep 435)
Brown, J. L., Jr., "A property of the generalized envelope," IRE Trans. on Circuit Theory, vol. CT-6, p. 325; September, 1959. (Dec 556)
Brown, W. M., "Some results on noise through circuits," IRE Trans. on Information Theory, vol. IT-5, pp. 217–227, May, 1959; IRE Trans. on Circuit Theory, vol. CT-6, pp. 217–227, May, 1959. (Mar 54) (Dec 556)
Burgess, T. E.: See Thomas, H. A.
Burnside, W., "Theory of Groups of Finite Order," Dover Publications, Inc., New York, N. Y.; 1955. (Sep 459)
Bussgang, J. J., and Middleton, D., "Optimum sequential detection of signals in noise," IRE Trans. on Information Theory, vol. IT-1, pp. 5–18; December, 1955. (Sep 505)
Bussgang, J. J., and Middleton, D., "Sequential Detection of Signals in Noise," Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rept. No. 175; August 31, 1955. (Sep 505)

C

Cahn, C. R.: See Bond, F. E.
Campbell, G. A., and Foster, R. M., "Fourier Integrals for Practical Applications," American Telephone and Telegraph Co.; 1942. (Apr 308)
Campbell, G. A., and Foster, R. M., "Fourier Integrals for Practical Applications," D. Van Nostrand Co., Inc., New York, N. Y.; 1947. (Apr 264)
Cauer, W., German Patent No. 892, 772; December 19, 1950. (Jun 328)
Cheatham, T. P., Jr., and Kohlenberg, A., "Analysis and Synthesis of Optical Systems," Phys. Res. Labs., Boston University, Boston, Mass., Tech. Note 84, pt. 1; March, 1952. (Jun 387)
Cheatham, T. P., Jr., and Kohlenberg, A., "Optical filters–their equivalence to and differences from electrical networks," 1954 IRE Convention Record, pp. 6–12. (Jun 387)
Church, R., "Tables of irreducible polynomials for the first four prime moduli," *Ann. Math.*, vol. 36, pp. 198–209; January, 1935. (Dec 525)
Cocke, J., "Lossless symbol coding with nonprimes," IRE Trans. on Information Theory, vol. IT-5, pp. 33–34; March, 1959. (Mar 15) (Dec 551)
Collar, A. R.: See Frazer, R. A.
Cook, C. E., "Modification of pulse compression waveforms," *Proc. 1958 Natl. Electronics Conf.*, pp. 1058–1067. (Jun 328)
Copson, E. T., "An Introduction to the Theory of a Complex Variable," Clarendon Press, Oxford, Eng.; 1935. (Apr 264)
Copson, E. T., "An Introduction to the Theory of Functions of a Complex Variable," Oxford University Press, New York, N. Y.; 1934. (Sep 477)
Coxeter, H. S. M., "Regular Polytopes," Methuen & Co., Ltd., London, Eng.; 1948. (Dec 516)
Cramer, H., "Mathematical Methods of Statistics," Princeton University Press, Princeton, N. J.; 1946. (Apr 264) (Jun 335)
Cutrona, L. J., *et al.*, "Optical data processing and filtering systems," IRE Trans. on Information Theory, vol. IT-6, p. 386–400; June, 1960. (Jun 328)

D

Darling, D. A., and Seigert, A. J. F., "A systematic approach to a class of problems in the theory of noise and other random phenomena—I," IRE Trans. on Information Theory, vol. IT-3, pp. 32–37; March, 1957. (Dec 539)
Darlington, S., U. S. Patent No. 2, 678, 997; May 18, 1954. (Jun 328, 385)
Davenport, W. B., Jr., "Signal-to-noise ratios in band-pass limiters," *J. Appl. Phys.*, vol. 24, pp. 720–727; June, 1953. (Jun 385)
Davenport, W. B., Jr., and Root, W. L., "An Introduction to Random Signals and Noise," McGraw-Hill Book Co., Inc., New York, N. Y.; 1958. (Jun 328, 349, 412)
Davie, I. L.: See Woodward, P. M.
Debart, H. P., "A new kind of matched filter," IRE Trans. on Circuit Theory, vol. CT-6, pp. 14–29; May, 1959. (Jun 349)
De Lano, R. H., "Reduction of Noise Fluctuation by Integration," Hughes Aircraft Co., Tech. Memo. No. 182; October, 1947. (Apr 266)
Delbruck, M.: See Golomb, S. W.

Den Hartog, H., and Muller, F. A., "Optimum instrument response for discrimination against spontaneous fluctuations," *Physica*, vol. 13, p. 571; 1947. (Jun 349)

"Development and Use of the Microband Lock-In Amplifier," Georgia Inst. Tech., Atlanta, NDRC, Div. 14, Rept. No. 592; September, 1945. (Apr 144, 266)

Dicke, R. H., "The measurement of thermal radiation at microwave frequencies," *Rev. Sci. Instr.*, vol. 17, pp. 268–275; July, 1946. (Sep 422)

Dicke, R. H., U. S. Patent No. 2, 624, 876; January 6, 1953. (Jun 328)

Dickson, L. E., "Introduction to the Theory of Numbers," Dover Publications, Inc., New York, N. Y.; 1957. (Mar 15)

Doob, L., "Stochastic Processes," John Wiley and Sons, Inc., New York, N. Y.; 1953. (Sep 444, 500) (Dec 540, 556)

Drenick, R.: See Balakrishnan, A. V.

Dugundji, J., "Envelopes and pre-envelopes of real waveforms," IRE Trans. on Information Theory, vol. IT-4, pp. 53–57; March, 1958. (Jun 330, 385) (Dec 556)

Dugundji, J., and Rice, S. O., "Mathematical analysis of random noise," in "Selected Papers on Noise and Stochastic Processes," Nelson Wax, Ed., Dover Publications, New York, N. Y.; 1954. (Jun 380) (Sep 485)

Duncan, W. J.: See Frazer, R. A.

Dwork, B. M., "Detection of a pulse superimposed on fluctuation noise," Proc. IRE, vol. 38, pp. 771–774; July, 1950. (Jun 329, 349)

Dwork, B. M., and Heller, R. M., "Results of a geometric approach to the theory and construction of nonbinary, multiple error and failure correcting codes," 1959 IRE National Convention Record, pt. 4, pp. 123–129. (Dec 523, 551)

E

Eaton, T. T.: See Wolff, I.

Elias, P., "Error free coding," IRE Trans. on Information Theory, vol. IT-4, pp. 29–37; September, 1954. (Sep 459)

Elias, P., "Optics and communication theory," *J. Opt. Soc. Am.*, vol. 43, pp. 229–232; April, 1953. (Jun 387)

Elias, P., Grey, D., and Robinson, D., "Fourier treatment of optical processes," *J. Opt. Soc. Am.*, vol. 42, pp. 127–134; February, 1952. (Jun 387)

Elspas, B., "A Radar System Based on Statistical Estimation and Resolution Considerations," D.Sc. dissertation, Stanford University, Stanford, Calif.; 1955. (Jun 385)

Elspas, B., "The theory of autonomous linear sequential networks," IRE Trans. on Circuit Theory, vol. CT-6, pp. 45–60; March, 1959. (Mar 14) (Jun 329) (Sep 503)

Elspas, B.: See Abramson, N.

Emslie, A. G., "Coherent Integration," Radiation Lab., Rept. No. 103. (Apr 143)

Erdelyi, A., *et al.*, "Higher Transcendental Functions," McGraw-Hill Book Co., Inc., New York, N. Y.; 1953. (Dec 543, 544)

F

Fano, R. M., "Communication in the Presence of Additive Gaussian Noise," in "Communication Theory," W. Jackson, Ed., Academic Press, Inc., New York, N. Y.; 1953. (Jun 329)

Fano, R. M., "On Matched-Filter Detection in the Presence of Multipath Propagation," Mass. Inst. Tech., Cambridge, unpublished paper; 1956. (Jun 329)

Faulkes, H. O., "Theorems of Kakeya and Polya on power sums," *Math. Z.*, vol. 65, pp. 345–352; 1956. (Sep 462)

Feller, W., "An Introduction to Probability Theory and Its Applications," John Wiley and Sons, Inc., New York, N. Y.; 1st ed. 1950, 2nd ed. 1957. (Sep 438, 485)

Fine, N. J., "Classes of periodic sequences," *Illinois J. Math.*, vol. 2, pp. 285–302; June, 1958. (Sep 477)

Fink, D. G., "Television Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., 2nd ed.; 1952. (Sep 504)

Fire, P., "A Class of Multiple-Error Correcting Binary Codes for Non-Independent Errors," Electronics Labs., Stanford University, Stanford, Calif., Tech. Rept. No. 55; April, 1959. (Mar 18)

Fire, P., "A Class of Multiple-Error-Correcting Binary Codes for Non-Independent Errors," Sylvania Reconnaissance Sys. Lab., Mountain View, Calif., Rept. No. RSL-E-2; March, 1959. (Mar 15)

Fisher, R. A., "The general sampling distribution of the multiple correlation coefficient," *Proc. Roy. Soc.* (*London*), Ser. A, vol. 81, pp. 654–672; December, 1928. (Apr 265)

Flejshman, B. S., "On the optimal detector with a log I_0 characteristic for the detection of a weak signal in the presence of noise," *Radiotekhnika i Elektronika*, vol. 2, pp. 726–734; June, 1957 (In Russian.) (Sep 505)

Fletcher, A., Miller, J. C. P., and Rosenhead, L., "An Index o Mathematical Tables," McGraw-Hill Book Company, Inc. New York, N. Y.; 1946. (Apr 264)

Fontaine, A. B., and Peterson, W. W., "Group code equivalence an optimum codes," IRE Trans. on Information Theory, vol IT-5, pp. 60–70; May, 1959. (Sep 466)

Fortet, R. M., "Average spectrum of a periodic series of identica pulses randomly displaced and distorted," *Elec. Commun.*, vol. 31 pp. 283–287; 1954. (Sep 444)

Fortet, R., "Random functions from a Poisson process," *Proc 2nd Berkeley Symp. on Math. Stat. and Prob. Theory*, J. Neyman Ed., University of California Press, Berkeley and Los Angeles pp. 375–385; 1951. (Sep 489)

Fortet, R.: See Blanc-Lapierre, A.

Foster, R. M.: See Campbell, G. A.

Fox, W. C.: See Petersen, W. W.

Franklin, G., "Linear filtering of sampled-data," 1955 IRE Con vention Record, pt. 4, pp. 119–128. (Sep 484)

Fraser, D. A. S., "Nonparametric Methods in Statistics," John Wiley and Sons, Inc., New York, N. Y.; 1957. (Jun 337)

Frazer, R. A., Duncan, W. J., and Collar, A. R., "Elementary Matrices," Cambridge University Press, Cambridge, Eng.; 1955 (Sep 428)

Friedland, B., "Theory of Time-Varying Sampled-Data Systems," Electronics Res. Lab., Columbia University, New York, N. Y. Tech. Rept. T-19/B; April, 1957. (Jun 366)

Fry, T. C., "Probability and its Engineering Uses," D. Van Nostran Co., Inc., New York, N. Y.; 1928. (Apr 264)

Fuller, A. T.: See Wonham, W. M.

G

Gabor, D., "Communication theory and physics," *Phil. Mag* vol. 41, pp. 1161–87; 1950. (Sep 435)

Gabor, D., "Lectures on Communication Theory," Res. Lab. o Electronics, Mass. Inst. Tech., Cambridge, Tech. Rept. No. 238 April 3, 1952. (Sep 435)

Gabor, D., "Theory of communications," *J. IEE*, vol. 93, pt. 3 pp. 429–457; November, 1956. (Jun 329, 330, 348, 385)

Gabor, D., and Lerner, R. M., "The representation of signals," IRE Trans. on Information Theory, vol. IT-5, pp. 197–216 May, 1959; Also IRE Trans. on Circuit Theory, vol. CT-6 pp. 197–216; May, 1959. (Jun 374)

Galejs, J., "Enhancement of pulse train signals by comb filters," IRE Trans. on Information Theory, vol. IT-4, pp. 114–125 September, 1958. (Jun 329)

Galejs, J., "Signal-to-noise ratios in smooth limiters," IRE Trans on Information Theory, vol. IT-5, pp. 79–85; June, 1959 (Mar 52)

Gamash, V. A., "Quantization of signals with nonuniform steps," *Electrosvyaz*, vol. 10, pp. 11–13; October, 1957. (Mar 8)

George, T. S., "Fluctuations of ground clutter return in airborn radar equipment," *J. IEE*, vol. 99, pt. 4, pp. 92–99; April, 1952 (Jun 329, 349)

Germond, H. H., and Hastings, C., "Preliminary Report on Scatte Bombing," NDRC Rept. No. 34; September, 1942. (Apr 266

Germond, H. H., and Hastings, C., "Scatter Bombing of a Circula Target," Appl. Math. Panel, Rept. No. 102 R, BRG Rept No. 120, AMG-C Rept. No. 302; May, 1944. (Apr 266)

Gilbert, E. N., "A comparison of signaling alphabets," *Bell Sys Tech. J.*, vol. 31, pp. 504–522; May, 1952. (Sep 466) (Dec 516

Gilbert, E. N., and Moore, E. F., "Variable-length binary encodings, *Bell Sys. Tech. J.*, vol. 38, pp. 933–967; July, 1959. (Sep 477

Girshick, M.: See Blackwell, D.

Gnedenko, B. V., and Kolmogorov, A. N., "Limit Distributions fo Sums of Independent Random Variables," translated by K. I Chung, Addison-Wesley Publishing Co., Cambridge, Mass 1954. (Sep 489)

Golay, M. J. E., "Notes on digital coding," Proc. IRE, vol. 37 p. 657; June, 1949. (Mar 13) (Sep 447, 459)

Golay, M. J. E., "Notes on the penny-weighing problem, lossle symbol coding with nonprimes, etc.," IRE Trans. on Informa tion Theory, vol. IT-4, pp. 103–109; September, 1958. (Dec 55

Goldman, S., "Information Theory," Prentice-Hall, Inc., New Yor N. Y.; 1953. (Jun 348)

Goldman, S., "Some fundamental considerations concerning nois reduction and range in radar and communication," Proc. IR vol. 36, pp. 584–594; May, 1948. (Apr 265)

oldstein, M. H., Jr.: See Booton, R. C., Jr.
oldstein, S., "On diffusion by discontinuous movements and on the telegraph equation," *Quart. J. Mech. and Appl. Math.*, vol. 4, pp. 129–156; 1951. (Dec 541)
oldstein, S. J., "A comparison of two radiometer circuits," Proc. IRE, vol. 43, pp. 1663–1666; November, 1955. (Sep 422)
olomb, S. W., "Sequences with the Cycle-and-Add Property," Jet Propulsion Lab., Calif. Inst. Tech., Pasadena, Section Rept. 8-573; December 19, 1957. (Jun 329)
olomb, S. W., "Sequences with Randomness Properties," Glenn L. Martin Co., Baltimore, Md., Internal Rept.; June 14, 1955. (Mar 18) (Jun 329) (Dec 523)
olomb, S. W., Gordon, B., and Welch, L. R., "Comma-free codes," *Can. J. Math.*, vol. 10, pp. 202–209; 1958. (Sep 477)
olomb, S. W., and Welch, L. R., "Nonlinear Shift-Register Sequences," Jet Propulsion Lab., Calif. Inst. Tech., Pasadena, Memo. 20–149; October 25, 1957. (Jun 329)
olomb, S. W., Welch, L. R., and Delbruck, M., "Construction and properties of comma-free codes," *Biol. Med. Danske Vid. Selsk.*, vol. 23, pp. 1–34; 1958. (Sep 477)
ordon, B.: See Golomb, S. W.
oudsmit, S. A., "The Comparison Between Signal and Noise," Rad. Lab., Mass. Inst. Tech., Lexington, Rept. No. 43–20; January, 1943. (Apr 143, 265)
oudsmit, S. A., "Statistics of Circuit Noise," Rad. Lab., Mass. Inst. Tech., Lexington, Rept. No. 43–20; January, 1943. (Apr 143, 265)
ray, A., Mathews, G. B., and Macrobert, T. M., "A Treatise on Bessel Functions," The Macmillan Company, New York, N. Y.; 1931. (Apr 264)
reen, J. H., Jr., and San Soucie, R. L., "An error-correcting encoder and decoder of high efficiency," Proc. IRE, vol. 46, pp. 1741–1744; October, 1958. (Mar 16) (Sep 465)
reen, P. E., Jr.: See Price, R.
rey, D.: See Elias, P.
uillemin, E. A., "The Mathematics of Circuit Analysis," John Wiley and Sons, Inc., New York, N. Y.; 1949. (Jun 364, 412)
uillemin, E. A., "Synthesis of Passive Networks," John Wiley and Sons, Inc., New York, N. Y.; 1957. (Jun 329, 372)

H

all, M., "Projective Planes and Related Topics," Calif. Inst. Tech., Pasadena; April, 1954. (Sep 459)
amming, R. W., "Error detecting and error correcting codes," *Bell Sys. Tech. J.*, vol. 29, pp. 147–160; April, 1950. (Mar 16) (Sep 444, 459, 503) (Dec 523, 557)
astings, C.: See Germond, H. H.
eatley, A. H., "A short table of the Toronto functions," *Trans. Roy. Soc. Canada*, vol. 37, pp. 13–29; 1943. (Apr 266)
eller, R. M.: See Dwork, B. M.
elstrom, C. W., The resolution of signals in white, Gaussian noise," Proc. IRE, vol. 43, pp. 1111–1118; September, 1955. (Jun 329, 367)
elstrom, C. W., "Single-error correcting codes for nonbinary symmetric channels," IRE Trans. on Information Theory, to be published. (Dec 551)
erbstreit, J. W., "Cosmic Noise," NBS, Rept. No. CRPL-4M-2; February, 1947. (Apr 143)
erlin, M. A.: See Allis, W. P.
erold, E. W., "Signal-to-noise ratio of UHF receivers," *RCA Rev.*, vol. 6, pp. 302–331; January, 1942. (Apr 143, 264)
erold, E. W., and Malter, L., "Some aspects of radio reception at ultra-high frequency," Proc. IRE, vol. 31, pp. 501–510; September, 1943. (Apr 265)
opper, A. L.: See Pierce, J. R.
uffman, D. A., "The synthesis of linear sequential coding networks," *3rd London Symp. on Information Theory*, London, Eng., Butterworth's Scientific Publications, London, Eng., Colin Cherry, Ed., pp. 77–95; 1956. (Dec 523)

J

affe, R., and Rechtin, E., "Design and performance of phase locked circuits capable of near optimum performance over a wide range of input signal and noise levels," IRE Trans. on Information Theory, vol. IT-1, pp. 66–76; March, 1955. (Sep 425)
ames, H. M., Nichols, N. B., and Phillips, R. S., "Theory of Servomechanisms," McGraw-Hill Book Co., Inc., New York, N. Y.; 1947. (Dec 539)
anos, W. A., "Optimal Filtering of Periodic Pulse-Modulated Time Series," Ph.D. dissertation, University of California, Berkeley; 1958. (Sep 484)
Jarema, E. D.: See Webb, H. D.
Jeans, J. H., "Kinetic Theory of Gases," Cambridge University Press, Cambridge, Eng.; 1940. (Dec 541)
Jordan, K. L., Jr., Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quart. Progr. Rept. No. 52; January, 1959. (Mar 53)
Jordan, W. H., "Action of Linear Detector on Signals in the Presence of Noise," Rad. Lab., Mass. Inst. Tech., Lexington, Rept. No. 61–23; July, 1943. (Apr 266)

K

Kac, M., and Siegert, A. J. F., "On the theory of noise in radio receivers with square law detectors," *J. Appl. Phys.*, vol. 18, pp. 383–397; April, 1947. (Apr 265)
Kailath, T., "Correlation detection of signals perturbed by a random channel," IRE Trans. on Information Theory, vol. IT-6, pp. 361–367; June, 1960. (Jun 329)
Kailath, T., "Sampling Models for Linear Time-Variant Filters," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Tech. Rept. No. 352; May, 1959. (Jun 357, 361, 412)
Kailath, T., "Solution of an integral equation occurring in multipath communication problems," IRE Trans. on Information Theory, vol. IT-6, p. 412; June, 1960. (Jun 357, 412)
Kanchisa, M.: See Nakagami, M.
Karlin, S.: See Bradt, R.
Keilson, J., Mermin, N. D., and Bello, P., "A theorem on cross correlation between noisy channels," IRE Trans. on Information Theory, vol. IT-5, pp. 77–79; June, 1959. (Mar 54)
Kelley, E. J.: See Reed, I. S.
Kendall, M. G., "The Advanced Theory of Statistics," Chas. Griffin and Company, Ltd., London, Eng., vols. 1 and 2; 1947. (Apr 264)
Kharkevitch, A. A., "To the theory of perfect receiver," *Electrosviaz*, vol. 10, pp. 28–34, 1956; Translation No. R-1956, U. S. Dept. of Commerce, Office of Technical Services, Washington, D. C.
Kohlenberg, A.: See Cheatham, T. P., Jr.
Kolmogrov, A., "Analytische Methoden in der Wahrscheinlichkeitsrechnung," *Math. Ann.*, vol. 104, pt. 2, pp. 415–458; 1931. (Dec 540)
Kolmogrov, A. N.: See Gnedenko, B. V.
Kramer, H. P., "Best finite linear approximation to a second-order stochastic process," *Notices Am. Math. Soc.*, vol. 5, p. 201; 1958. (Mar 52)
Kramer, H. P. and Mathews, M. V., "A linear coding for transmitting a set of correlated signals," IRE Trans. on Information Theory, vol. IT-2, pp. 41–46; September, 1956. (Mar 52)
Kronert, R., "Impulsverdichtung," *Nachrichtentech.*, vol. 7, pp. 148–152, 162, April, 1957; pp. 305–308, July, 1957. (Jun 329)
Kuebler, R. R., Jr.: See Bose, R. C.

L

Laha, R. G., 'On a characterization of stable law with finite expectation," *Ann. Math. Stat.*, vol. 27, pp. 187–195; March, 1956. (Sep 500)
Lampard, D. G., and Barrett, J. F., "An expansion for some second-order probability distributions and its application to noise problems," IRE Trans. on Information Theory, vol. IT-1, pp. 10–16; March, 1955. (Sep 502)
Landon, V. D., "The distribution of amplitude with time in fluctuation noise," Proc. IRE, vol. 30, pp. 425–429; September, 1942. (Apr 265)
Laning, J. H., and Battin, R. H., "Random Processes in Automatic Control," McGraw-Hill Book Co., Inc., New York, N. Y.; 1956. (Mar 54)
Lause, H. W., Peeler, G. D. M., and Randall, C. R., "Video Mixing and Minimum Detectable Signal in the SPS-3 Radar," Nav. Res. Lab., Rept. R-3123; September, 1947. (Apr 266)
Lawson, J. L., and Uhlenbeck, G. E., "Threshold Signals," Mass. Inst. Tech., Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 24; 1950. (Apr 144, 264) (Jun 349)
Lee, C. Y., Some properties of nonbinary error-correcting codes," IRE Trans. on Information Theory, vol. 4, pp. 77–82; June, 1958. (Dec 523, 551)
Lees, A. B., Interpolation and extrapolation of sampled-data," IRE Trans. on Information Theory, vol. IT-2, pp. 12–17; March, 1956. (Sep 484)

Lehan, F. W. and Parks, R. J., "Optimum demodulation," 1953 IRE Convention Record, pt. 8, pp. 101–103. (Sep 420)

Lerner, R. M., "A matched filter detection system for doppler-shifted signals," IRE Trans. on Information Theory, vol. IT-6, pp. 373–385; June, 1960. (Jun 329)

Lerner, R. M., "Delay lines for use in signal detection," IRE Trans. on Ultrasonics Engineering, to be published. (Jun 382)

Lerner, R. M.: See Gabor, D.

Lerner, R. M.: See Manasse, R.

Lerner, R. M., "Signals with uniform ambiguity functions," 1958 IRE National Convention Record, pt. 4, pp. 27–36. (Jun 329, 348, 385)

Lerner, R. M., Reiffen, B., and Sherman, H., "Delay-line specifications for matched-filter communications systems," IRE Trans. on Component Parts, vol. CP-6, pp. 263–268; December, 1959. (Jun 382)

Levin, M. J., "Optimum estimation of impulse response in the presence of noise," 1959 IRE National Convention Record, pt. 4, pp. 174–181. (Jun 329)

Levine, D.: See Staley, J.

Levine, A. and McGhee, R. B., "Cumulative distribution functions for a sinusoid plus Gaussian noise," IRE Trans. on Information Theory, vol. IT-5, pp. 90–91; June, 1959. (Sep 502)

Levinson, N., "The Wiener RMS (root mean square) error criterion in filter design and prediction," *J. Math. Phys.*, vol. 25, pp. 261–278; January, 1947. (Jun 366)

Lindley, D., "On a measure of the information provided by an experiment," *Ann. Math. Stat.*, vol. 27, pp. 986–1005; 1956. (Dec 530)

Lloyd, S. P., "Binary block coding," *Bell Sys. Tech. J.*, vol. 36, pp. 517–535; March, 1957. (Dec 523)

Loeve, M., "Probability Theory," D. Van Nostrand Co., Inc., New York, N. Y.; 1955. (Mar 53) (Sep 500)

Lowan, A. N., "Tables of Probability Functions," NBS, Washington, D. C., vol. 2; 1942. (Apr 264)

Lytle, D. W., "Experimental Study of Tapped Delay-Line Filters," Electronics Labs., Stanford University, Stanford, Calif., Tech. Rept. No. 361–3; July 30, 1956. (Jun 329)

Lytle, D. W., "On the Properties of Matched Filters," Electronics Labs., Stanford University, Stanford, Calif., Tech. Rept. No. 17; June 10, 1957. (Jun 329)

M

MacDonald, J. E., "Design methods for maximum minimum-distance error-correcting codes," *IBM J. Res. Dev.*, vol. 4, pp. 43–57; January, 1960. (Dec 552)

McFadden, J. A., "The probability density of the output of an RC filter when the input is a binary random process," IRE Trans. on Information Theory, vol. IT-5, pp. 174–178; December, 1959. (Dec 540)

McFadden, J. A., "The probability density of the output of a filter when the input is a random telegraphic signal," 1959 IRE National Convention Record, pt. 4, pp. 164–169. (Dec 541)

McFadden, J. A., "The probability density of the output of a filter when the input is a random telegraphic signal: differential equation method," IRE Trans. on Circuit Theory, vol. CT-6, pp. 228–233; May, 1959. (Dec 540)

MacKay, D. M., "Quantal aspects of scientific information," IRE Trans. on Information Theory, vol. IT-1, pp. 60–80; February, 1953. (Sep 435)

MacLane, S.: See Birkoff, G.

Macrobert, T. M.: See Gray, A.

Malter, L.: See Herold, E. W.

Manasse, R., "Parameter Estimation Theory and Some Applications of the Theory to Radar Measurements," Mitre Corp., Tech. Rept., to be published. (Jun 340)

Manasse, R., "Range and Velocity Accuracy from Radar Measurements," Lincoln Lab., Mass. Inst. Tech., Lexington, Group Rept. 312–26; February, 1955. (Jun 330)

Manasse, R., Price, R., and Lerner, R. M., "Loss of signal detectability in band-pass limiters," IRE Trans. on Information Theory, vol. IT-4, pp. 34–38; March, 1953. (Jun 385)

Marcum, J. I., "A Statistical Theory of Target Dectection by Pulsed Radar," Project RAND, Douglas Aircraft Co., Inc., Rept. No. RA-15061; December 1, 1947 (confidential). (Apr 265)

Marcum, J. I., "A Statistical Theory of Target Detection by Pulsed Radar," RAND Corp., Santa Monica, Calif., Repts. RM-753 and RM-754; July, 1948, and December, 1947. (Apr 308) (Jun 329)

Marcum, J. I., "A Statistical Theory of Target Detection by Pulsed Radar—Mathematical Appendix," Project RAND, Dougl Aircraft Co., Inc., Rept. No. RA-15062; to be published. (Apr 14

Margenau, H. and Murphy, G. M., "The Mathematics of Physi and Chemistry," D. Van Nostrand Co., Inc., New York, N. Y 1943. (Jun 354)

Martin, R. W.: See Siegert, A. J. F.

Marsh, R. W.: See Sloan, R. W.

Marsh, R. W., "Table of Irreducible Polynomials Over GF(Through Degree 19," NSA, Washington, D. C.; October 2 1957. (Mar 18) (Sep 503)

Masonson, M., 'Binary transmissions through noise and fading 1957 IRE National Convention Record, pt. 2, pp. 69–8 (Dec 558)

Mathews, G. B.: See Gray, A.

Mathews, M. V.: See Kramer, H. P.

McAffee, W. S.: See Webb, H. D.

McGhee, R. B.: See Levine, A.

Meggitt, J. E., "Error-correcting codes for correcting bursts errors," to be published. (Sep 503)

Mermin, N. D.: See Keilson, J.

Metzler, W. H.: See Muir, T.

Middleton, D., "An Introduction to Statistical Communicatio Theory," McGraw-Hill Book Co., Inc., New York, N. Y.; 196 (Jun 349) (Sep 505) (Dec 549)

Middleton, D., "On the detection of stochastic signals in additiv normal noise—Part I," IRE Trans. on Information Theor vol. IT-3, pp. 86–121; June, 1957. (Jun 350, 412)

Middleton, D., "On new classes of matched filters and generalizatio of the matched filter concept," IRE Trans. on Informatio Theory, vol. IT-6, pp. 349–360; June, 1960. (Jun 412)

Middleton, D.: See Bussgang, J. J.; Van Meter, D.; Van Vleck, J. I

Middleton, D., "Some general results in the theory of noise throug nonlinear devices," *Quart. Appl. Math.*, vol. 5; January, 194 (Apr 265)

Middleton, D., "Statistical criteria for the detection of pulse carriers in noise (pts. I and II)," *J. Appl. Phys.*, vol. 24, pp. 371 378, 379–391; April, 1953. (Sep 504)

Middleton, D., and Van Meter, D., "Detection and extraction signals in noise from the point of view of statistical decisio theory," *J. Soc. Indus. Appl. Math.*, vol. 3, pp. 192–253, Decembe 1955; vol. 4, pp. 86–119, June, 1956. (Mar 4) (Jun 329, 350, 36

Miller, J. C. P.: See Fletcher, A.

Miller, K. S., Bernstein, R. I., and Blumenson, L. E., "Generalize Rayleigh processes," *Quart. Appl. Math.*, vol. 16; July, 195 (Sep 502)

Milne, W. E., "Numerical Calculus," Princeton University Pres Princeton, N. J., 3rd ed.; 1949. (Sep 427)

Mittra, R., "On the Solution of a Class of Wiener-Hopf Integr Equations in Finite and Infinite Ranges," Antenna Lab., Un versity of Illinois, Urbana, Tech. Rept. No. 37; 1959. (Sep 48

Moore, E. F.: See Gilbert, E. N.

Moore, J. R., "Computation Search Time, Power Requiremen Range, and Other Parameters for Radar Systems," Navy Dept BuAer., Design Res. Div., Rept. No. 1022; May, 1947. (Apr 14

Muir, T., and Metzler, W. H., "A Treatise on the Theory of Dete minants," Longmans, Green, and Co., New York, N. Y., ch. 2 1933. (Sep 461)

Muller, F. A., 'Communication in the Presence of Additive Gaussi Noise," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridg Tech. Rept. 244; May 27, 1953. (Jun 329)

Muller, F. A.: See Den Hartog, H.

Murphy, G. M.: See Margenau, H.

N

Nakagami, M., Tanaka, K., and Kanchisa, M., "The *m*-distributi as the general formula of intensity distribution of rapid fading *Mem. Fac. Engrg.*, Kobe University, Japan; March, 1957. (Sep 50

Newman, M. H. A., "Elements of the Topology of Plain Sets Points," Cambridge University Press, Cambridge, Eng.; 19 (Dec 524)

Nichols, N. B.: See James, H. M.

North, D. O., "Analysis of Factors which Determine Signal-to-No Discrimination in Pulsed Carrier Systems," RCA Labs., Princet N. J., Rept. No. PTR-6C; June, 1943. (Apr 266) (Jun 361)

North, D. O., "Analysis of the Factors which Determine Signal-t Noise Discrimination in Radar," RCA Labs., Princeton, N. Rept. No. PTR-6C; June, 1943. (Apr 143) (Jun 329, 349)

North, D. O., "The Absolute Sensitivity of Radio Receiver *RCA Rev.*, vol. 6, pp. 332–343; January, 1942. (Apr 143)

rton, K. A., and Omberg, A. C., "The Maximum Range of a Radar Set," Proc. IRE, vol. 35, pp. 4–24; January, 1947. (Apr 43, 264)

O

nberg, A. C.: See Norton, K. A.

Meara, T. R., "The Synthesis of 'Band-Pass,' All-Pass Time-Delay Networks with Graphical Approximation Techniques," Hughes Res. Labs., Culver City, Calif., Res. Rept. 114; June, 1959. Jun 329)

Neill, E., "Selected Topics in Optics and Communication Theory," Boston University, Boston, Mass., Tech. Note 133; October, 1957. Jun 387)

Neill, E., "Spatial filtering in optics," IRE Trans. on Information Theory, vol. IT-2, pp. 56–65; June, 1956. (Jun 387)

Neill, E., "The Analysis and Synthesis of Linear Coherent Optical System," Phys. Res. Labs., Boston University, Boston, Mass., Tech. Note 122; September, 1955. (Jun 387)

P

ige, L. J., "A note on the Mathieu groups," *Can. J. Math.*, vol. , pp. 15–18; January, 1956. (Sep 459)

ley, R. E. A. C., "On orthogonal matrices," *J. Math. and Phys.*, vol. 12, pp. 311–320; 1933. (Sep 447)

rks, R. J.: See Lehan, F. W.

rzen, E., "Asymptotically Stationary Stochastic Processes," unpublished). (Sep 444)

arson, K., "Tables of the Incomplete Gamma Function," Cambridge University Press, Cambridge, Eng., (published by *Biometrika*); 1946. (Apr 264)

eler, G. D. M.: See Lause, H. W.

terson, W. W.: See Fontaine, A. B.

terson, W. W., and Birdsall, T. G., "The Theory of Signal Detectability," Electronics Defense Group, University of Michigan, Ann Arbor, Tech. Rept. No. 13; 1953. (Jun 409)

terson, W. W., Birdsall, T. G., and Fox, W. C., "Theory of signal detectability," IRE Trans. on Information Theory, vol. IT-4, pp. 171–212; September, 1954. (Jun 329, 385) (Sep 504)

trovsky, I. G., "Lectures on Partial Differential Equations," Interscience Publishers, Inc., New York, N. Y., 1954. (Dec 541)

illips, R. S.: See James, H. M.

rce, J. N., "Theoretical diversity improvement in frequency-shift keying," Proc. IRE, vol. 46, pp. 903–910; May, 1958. Jun 329)

rce, J. R., and Hopper, A. L., "Nonsynchronous time division with holding and with random sampling," Proc. IRE, vol. 40, pp. 1079–1088; September, 1950. (Sep 444)

cher, T. S.: See Root, W. L.

otkin, M., "Binary Codes with Specified Minimum Distance," M.S. thesis, Moore School of Elec. Engrg., University of Pennsylvania, Philadelphia; June, 1952. (Dec 524)

ritsky, H., "The Distribution of Points on a Sphere and Its Application to a Problem in Communication Theory," General Electric Res. Lab., Schenectady, N. Y., Rept. No. 57-RL-1765; August, 1957. (Dec 516)

nge, E., "Cyclic Error-Correcting Codes in Two Symbols," AFCRC TN-57-103, ASTIA Doc. No. AD 133749; September, 1957. (Sep 459)

nge, E., "Some Cyclic Error-Correcting Codes with Simple Decoding Algorithms," AFCRC-TN-58-156, ASTIA Doc. No. AD 152386; April, 1958. (Sep 459)

nge, E., "The Use of Coset Equivalence in the Analysis and Decoding of Group Codes," AFCRC-TR-59-164; June, 1959. Sep 459)

ce, R., "Error Probabilities for the Ideal Detection of Signals Perturbed by Scatter and Noise," Lincoln Lab., Mass. Inst. Tech., Lexington, Group Rept. No. 34–40; October 3, 1955. (Jun 329)

ce, R., "Optimum detection of random signals in noise, with application to scatter-multipath communication," IRE Trans. on Information Theory, vol. IT-2, pp. 125–135; December, 1956. (Jun 329, 350, 365)

ce, R.: See Manasse, R.

ce, R., and Green, P. E., Jr., "A communication technique for multipath channels," Proc. IRE, vol. 46, pp. 555–570; March, 1958. (Jun 329, 367, 385, 410)

gachev, V. S., "Determination of an optimum system using a general criterion," *Avtomat. i Telemekh.*, vol. 19, pp. 519–539; June, 1958 (in Russian); also *Automation and Remote Control*, same issue (in English). (Mar 4)

Pugachev, V. S., "Integral canonic representation of random functions and their application in deriving optimal linear systems," *Avtomat. i Telemekh.*, vol. 18, pp. 971–984, November, 1957 (in Russian); also *Automation and Remote Control*, same issue (in English). (Mar 6)

Pugachev, V. S., "Method of solving the basic integral equation of the statistical theory of optimum systems in closed form," *Prikladnaia Matemat. i Mekh.*, vol. 23, pp. 3–14; January, 1959. (Mar 6)

Pugachev, V. S., "Use of canonical expansions of random functions to derive optimal linear systems," *Avtomat. i Telemekh.*, vol. 17, pp. 489–499; June, 1956 (in Russian); also *Automation and Remote Control*, same issue (in English). (Mar 5)

Pugachev, V. S., "Theory of Random Functions and its Application to Automatic Control Problems," Technical Literature, Moscow, Russia; 1957. (In Russian.) (Mar 5)

R

"Radar range calculator," *Bell Labs. Rec.*, vol. 29, pp. 312–313; 1951. (Apr 144, 266)

"Radar System Analysis," Sperry Gyroscope Co., Rept. No. 5223–1109; June, 1948. (Apr 266)

The RAND Corp., "A Million Random Digits with 100,000 normal Deviates," Free Press, Glencoe, Ill.; 1955. (Sep 486)

Randall, C. R.: See Lause, H. W.

Ragazzini, J. R.: See Zadeh, L. A.

Ray-Chaudhuri, D. K.: See Bose, R. C.

Rechtin, E.: See Jaffe, R.

Reed, I. S., "A class of multiple-error-correcting codes and the decoding scheme," IRE Trans. on Information Theory, vol. IT-4, pp. 38–49; September, 1954. (Mar 16) (Jun 402) (Sep 459)

Reed, I. S., "On the use of Laguerre polynomials in treating the envelope and phase components of narrow-band Gaussian noise," IRE Trans. on Information Theory, vol. IT-5, pp. 102–105; September, 1959. (Sep 502)

Reed, I. S., Kelley, E. J., and Root, W. L., "The Detection of Radar Errors in Noise. Part II: The Accuracy of Radar Measurements," Lincoln Lab., Mass. Inst. Tech., Lexington, Tech. Rept. No. 159; July 19, 1957. (Jun 337)

Reed, I. S. and Solomon, G., "Polynomial code," to be published in *J. Soc. Ind. Appl. Math.* (Sep 459)

Reich, E. and Swerling, P., "The detection of a sine wave in Gaussian noise," *J. Appl. Phys.*, vol. 24, pp. 289–296; March, 1953. (Jun 329)

Reiffen, B.: See Lerner, R. M.

Reiger, S., "Error probabilities of binary data transmission in the presence of random noise," 1953 IRE Convention Record, pt. 8, pp. 72–79. (Jun 329) (Dec 551)

Reiger, S., "Error rates in data transmission," Proc. IRE, vol. 46, p. 919; May, 1958. (Dec 551)

Rhodes, J., "Analysis and synthesis of optical images," *Am. J. Phys.*, vol. 21, pp. 337–343; January, 1953. (Jun 387)

Rice, S. O., "Communication in the presence of noise—probability of error for two encoding schemes," *Bell Sys. Tech. J.*, vol. 29, pp. 60–69; January, 1950. (Sep 502) (Dec 516)

Rice, S. O., "Filtered thermal noise-fluctuation of energy as a function of interval length," *J. Acoust. Soc. Am.*, vol. 14, pp. 216–227; 1943. (Apr 265)

Rice, S. O., "The mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, pp. 282–332, July, 1944; vol. 24, pp. 46–156, January, 1945. (Apr 144, 265) (Sep 502, 504) (Dec 540)

Rice, S. O., "The mathematical analysis of random noise," reprinted in the collection "Selected Papers on Noise and Stochastic Processes," N. Wax, Ed., Dover Publications, Inc., New York, N. Y., pp. 133–294; 1954. (Dec 549, 558)

Rice, S. O.: See Dugundji, J.

Rice, S. O., "Statistical properties of a sine wave plus random noise," *Bell Sys. Tech. J.*, vol. 27, pp. 109–157; January, 1948. (Apr 265) (Sep 502)

Riordan, J., "An Introduction to Combinatorial Analysis," John Wiley and Sons, Inc., New York, N. Y.; 1958. (Sep 461, 477)

Ridenour, L. N., "Radar System Engineering," Mass. Inst. Tech., Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 1; 1947. (Apr 144, 264)

Robinson, D.: See Elias, P.

Rochefort, J. S., "Matched filters for detecting pulsed signals in noise," 1954 IRE Convention Record, pt. 4, pp. 30–34. (Jun 329)

Root, W. L.: See Davenport, W. B.

Root, W. L.: See Reed, I. S.

Root, W. L. and Pitcher, T. S., "Some remarks on statistical detection," IRE Trans. on Information Theory," vol. IT-1, pp. 33–38; December, 1955. (Jun 329)

Rosenhead, L.: See Fletcher, A.

S

Sachs, G. E., "Multiple error correcting by means of parity checks," IRE Trans. on Information Theory, vol. IT-4, pp. 145–147; December, 1958. (Mar 17)

Sanders, R. W., "Communication efficiency of space telemetry systems," *Proc. Natl. Telemetry Conf.*, Denver, Colo.; May, 1959. (Jun 410)

San Soucie, R. L.: See Green, J. H., Jr.

Schmidt, E., "Zur Theorie der linearen und nicht linearen Integralgleichungen," *Math. Ann.*, vol. 63, pp. 433–476; 1907. (Mar 52)

Schoute, P. H., "Mehridimenzionale Geometre II. Teil, Die Polytope," Goschensche Verlaghandlung, Leipzig, Ger.; 1905. (Dec 516)

Scott, B. L., and Welch, L. R., "An Investigation of Iterative Boolean Sequences," Jet Propulsion Lab., Calif. Inst. Tech., Pasadena, Section Rept. 8–543; November 1, 1955. (Jun 329)

Shannon, C. E., "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379–423, 623–656; July/October, 1948. (Mar 22) (Dec 516)

Shannon, C. E., "A note on a partial ordering for communication channels," *Information and Control*, vol. 1, pp. 390–397; December, 1958. (Dec 531)

Shannon, C. E., "Certain results in coding theory for noisy channels," *Information and Control*, vol. 1, pp. 6–25; September, 1957. (Mar 22)

Shannon, C. E., "Communication in the presence of noise," Proc. IRE, vol. 37, pp. 10–21; January, 1949. (Jun 329, 373) (Dec 516)

Shannon, C. E., and Weaver, W., "The Mathematical Theory of Communication," University of Illinois Press, Urbana; 1949. (Sep 435, 489)

Shapiro, H. S., and Slotnick, D. L., "On the mathematical theory of error-correcting codes," *IBM J. Res. & Dev.*, vol. 3, pp. 25–34; January, 1959. (Dec 551)

Sherman, H.: See Lerner, R. M.

Siddiqui, M. M., "The inversion of the sample covariance matrix in a stationary auto-regressive process," *Ann. Math. Stat.*, vol. 129, pp. 585–588; June, 1958. (Sep 484)

Siebert, W. M., "A radar detection philosophy," IRE Trans. on Information Theory, vol. IT-2, pp. 204–221; September, 1956. (Jun 329, 348, 385, 400)

Siebert, W. M., "Studies of Woodward's Uncertainty Functions," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quart. Progr. Rept., pp. 90–94; April 15, 1958. (Jun 329, 348, 385)

Siegert, A. J. F.: See Darling, D. A.

Siegert, A. J. F.: See Kac, M.

Siegert, A. J. F., and Martin, F. W., "Fluctuations in Return Signals from Random Scatterers," Rad. Lab., Mass. Inst. Tech., Cambridge, Rept. No. 773, January, 1946. (Apr 266)

Silverman, R. A., "The theory of an isospectral family of random processes," IRE Trans. on Information Theory, vol. IT-6, pp. 485–490 September, 1960. (Dec 548)

Slepian, D., "A class of binary signaling alphabets," *Bell Sys. Tech. J.*, vol. 35, pp. 203–234; January, 1956. (Mar 13, 16) (Sep 459) (Dec 523, 524, 551)

Slepian, D., "Estimation of signal parameters in the presence of noise," IRE Trans. on Information Theory, vol. IT-3, pp. 68–89; March, 1954. (Mar 6)

Sloan, R. W., and March, R. W., "The Structure of Irreducible Polynomials Mod 2 Under a Cubic Transformation," private communication; July 7, 1953. (Jun 329)

Slotnick, D. L.: See Shapiro, H. S.

Smith, W. L., "Renewal theory and its ramifications," *J. Roy. Statistical Soc.*, ser. B, vol. 20, no. 2, pp. 243–284; 1958. (Sep 444)

Smithies, F., "Integral Equations," Cambridge University Press, Cambridge, Eng., 1958. (Mar 53)

Solomon, G.: See Reed, I. S.

Sommerville, D. M. Y., "An Introduction to the Geometry of N Dimensions," Dover Publications, Inc., New York, N. Y.; 1958. (Dec 516)

"Space Handbook: Astronautics and Its Applications, Staff Report of the Select Committee on Astronautics and Space Exploration," U. S. Govt. Printing Office, Washington, D. C., House Doc. No. 86, p. 181; 1959. (Sep 503)

Staley, J. and Levine, D., "Long-Range Search Radar for the Army Air Forces," Serial No. TSELR-113; January, 1947. (Apr 266)

Stern, T. E., "Information rates in photon channels and photo amplifiers," 1960 IRE International Convention Recor pt. 4, pp. 182–189. (Sep 435)

Stewart, J. L. and Westerfield, E. C., "A theory of active son detection," Proc. IRE, vol. 47, pp. 872–881; May, 1959. (Jun 34

Storer, J. E., and Turyn, R., "Optimum finite code groups," Pro IRE, vol. 46, p. 1649; September, 1958. (Jun 329)

Storer, J. E., and Turyn, R., "Optimum finite codes," *Proc. Sym on Recent Advances in Matched Filter Theory and Pulse Compressio Techniques*, Rome Air Dev. Command, ARDC, USAF, Griffi AFB, N. Y., RADC-TR-59-161; September, 1959. (Jun 40

Strutt, M. J. O., "Ultra and Extreme Short Wave Reception D. Van Nostrand Co., Inc., New York, N. Y.; 1947. (Apr 26

Stutt, C. A., "Regular Polyhedron Codes," General Electric Re Lab., Schenectady, N. Y., Rept. No. 59-RL-2202; March, 195 (Dec 516, 519, 520, 521)

Sussman, S., "A matched-filter communciation system for multipat channels," IRE Trans. on Information Theory, vol. IT- pp. 367–373; June, 1960. (Jun 329, 385)

Sutro, P. J., and Agathen, B., "The Theoretical Effect of Integratio on the Visibility of Weak Signals Through Noise," Harvar Radio Res. Lab., Cambridge, Mass., Rept. No. 411–77; Februar 1944. (Apr 266)

Swerling, P., "Optimum linear estimation for random process as the limit of estimates based on sampled data," 1958 WESCO Convention Record, pt. 4, pp. 158–163. (Sep 484)

Swerling, P.: See Reich, E.

Sydoriak, S., "The Effects of Video Mixing Ratio and Limitin on Signal Threshold Power," Nav. Res. Lab., Washingto D. C., NRL Rept. No. R-3008; February, 1946. (Apr 143, 26

Szego, G., "On the boundary value of an analytic function," *Mat Ann.*, vol. 84, pp. 232–244; 1921. (Sep 484)

T

Tanaka, K.: See Nakagami, M.

Thomas, J. B., "On the Statistical Design of Demodulation System for Signals in Additive Noise," Electronics Res. Lab., Stanfo University, Stanford, Calif., Tech. Rept. No. 88; August, 195 (Sep 424)

Thomas, H. A., and Burgess, R. E., "Survey of Existing Informatio and Data on Radio Noise and the Frequency Range 1–30 Mc/s. Dept. of Scientific and Industrial Res., His Majesty's Statione Office, London, Eng.; 1947. (Apr 267)

Thomas, H. B., Williams, T. R., Wolf, J., and Wong, E., "T demodulation of AM signals in noise," *Proc. 1959 Conventio on Military Electronics*, pp. 138–146. (Sep 424)

Titchmarsh, E. C., "Introduction to the Theory of Fourier Int grals," Oxford University Press, London, Eng.; 1937. (Sep 42

Titchmarsh, E. C., "The Theory of Functions," Oxford Universit Press, New York, N. Y., 2nd ed.; 1939. (Sep 477)

Titchmarsh, E. C., "The zeros of certain integral functions *Proc. London Math. Soc.*, vol. 25, pp. 283–302; May 14, 192 (Mar 51)

Titsworth, R. C., "Coherent detection by quasi-orthogonal squar wave pulse functions," IRE Trans. on Information Theor vol. IT-6, pp. 410–411; June, 1960. (Jun 329)

Titsworth, R. C., and Welch, L. R., "Modulation by Random a Pseudo-Random Sequences," Jet Propulsion Lab., Calif. Ins Tech., Pasadena, Prog. Rept. No. 20–397; June 5, 1959. (Jun 41

Tuller, W. G., "Theoretical limitations on rate of informati transmission," Proc. IRE, vol. 37, p. 4681; May, 1949. (Dec 51

Turin, G. L., "A Review of Correlation, Matched Filter, and Sign Coding Techniques, with Emphasis on Radar Applications Systems Dev. Labs., Hughes Aircraft Co., Los Angeles, Cali Tech. Memo. 559. (Jun 385)

Turin, G. L., "An introduction to matched filters," IRE Trans. Information Theory, vol. IT-6, pp. 311–329; June, 196 (Jun 385)

Turin, G. L., "Communication through noisy, random-multipa channels," 1956 IRE Convention Record, pt. 4, pp. 154–16 (Jun 329, 367)

Turin, G. L., "Error probabilities for binary symmetrical ide reception through nonselective slow fading and noise," Pr IRE, vol. 46, pp. 1603–1619; September, 1958. (Jun 329, 35

Turin, G. L., "On the estimation in the presence of noise of t impulse response of a random, linear filter," IRE Trans. Information Theory, vol. IT-3, pp. 5–10; March, 1957. (Jun 32

Turin, G. L., "Some computations of error rates for selectively fadi multipath channels," *Proc. 1959 NEC*, pp. 431–440. (Jun 329)

U

hlenbeck, G. E.: See Lawson, J. L.
hlenbeck, G. E., "Theory of Random Processes," Rad. Lab., Mass. Inst. Tech., Cambridge, Rept. No. 454; October, 1943. (Apr 143, 266)
lrich, W., "Nonbinary error correction codes," *Bell Sys. Tech. J.*, vol. 36, pp. 1341–1388; November, 1957. (Dec 511)
rkowitz, H., "Analysis and synthesis of delay line periodic filters," IRE Trans. on Circuit Theory, vol. CT-4, pp. 41–53; June, 1957. (Jun 329)
rkowitz, H., "Filters for detection of small radar signals in clutter," *J. Appl. Phys.*, vol. 24, pp. 1024–1031; August, 1953. (Jun 348, 349)

V

an der Waerden, B. L., "Modern Algebra," F. Ungar Publishing Co., New York, N. Y., vols. 1 and 2; 1949, 1950. (Sep 463) (Dec 551)
an Meter, D.: See Middleton, D.
an Meter, D., and Middleton, D., "Modern statistical approaches to reception in communication theory," IRE Trans. on Information Theory, vol. IT-4, pp. 119–145; September, 1954. (Mar 4)
an Vleck, J. H., and Middleton, D., "Theoretical comparison of the visual, aural, and meter reception of pulsed signals in the presence of noise," *J. Appl. Phys.*, vol. 17, pp. 940–971; November, 1946. (Apr 143, 265) (Jun 329, 349, 361)

W

agner, A., "Theorie der Boigkeit und der Hautigkeitsberteilung von Windstarke und Windrichtung," *Gerland's Beitrage zur Geophysik*, vol. 24, pp. 386–436; 1929. (Apr 265)
ald, A., "Statistical Decision Functions," John Wiley and Sons, Inc., New York, N. Y.; 1950. (Mar 4)
atson, G. N., "A Treatise on the Theory of Bessel Functions," The Macmillan Company, New York, N. Y.; 1944. (Apr 264)
ax, N., "On upper bounds for error-detecting and error-correcting codes of finite length," IRE Trans. on Information Theory, vol. IT-5, p. 168; December, 1959. (Dec 557)
eaver, W.: See Shannon, C. E.
ebb, H. D., McAfee, W. S., and Jarema, E. D., "CW Injection as a Means of Decreasing the Minimum Detectable Signal of a Radar Receiver," Camp Evans Signal Lab., Tech., Rept. No. T-32; December, 1944. (Confidential.) (Apr 143, 267)
elch, L. R.: See Golomb, S. W.
elch, L. R.: See Scott, B. L.
elch, L. R.: See Titsworth, R. C.
esterfield, E. C.: See Stewart, J. L.
iener, N., "Extrapolation, Interpolation and Smoothing of Stationary Time Series," John Wiley and Sons, Inc., New York, N. Y.; 1949. (Sep 484)
iener, N., "Nonlinear Problems in Random Theory," Technology Press, Mass. Inst. Tech., Cambridge, and John Wiley and Sons, Inc., New York, N. Y.; 1959. (Jun 403)
illiams, T. R.: See Thomas, J. B.
inkelbauer, K., "Experience in games of strategy and in statistical decision," *Trans. 1st Prague Conf. on Information Theory, Statistical Decision Functions and Random Processes*, Liblice, November 28–30, 1956, Czech. Acad. Sc., Prague, Czechoslovakia, pp. 297–354, 1957. (Mar 7)
ise, J., "The autocorrelation function and the spectral density function," *Biometrika*, vol. 42, pp. 151–159; 1955. (Sep 484)
Wold, H., "A Study in the Analysis of Stationary Time Series," Almquist and Wiksell, Stockholm, Sweden; 1938, 1954. (Sep 484)
Wolf, W.: See Born, M.
Wolf, J.: See Thomas, J. B.
Wolff, I., "Experimental Effect of Integration," RCA Labs., Princeton, N. J., Rept. No. PTR-76. (Apr 143)
Wolff, I., and Eaton, T. T., "An Experimental Investigation of the Factors Which Determine Signal Noise Discrimination in Pulse Radar Systems," RCA Labs., Princeton, N. J., Rept. No. PTR-7C; June, 1943. (Apr 266)
Wong, E.: See Thomas, J. B.
Wonham, W. M., "Transition probability densities of the smoothed random telegraph signal," *J. Electronics and Control*, vol. 6, pp. 376–384; 1959. (Dec 541)
Wonham, W. M., and Fuller, A. T., "Probability densities of the smoother random telegraph signal," *J. Electronics and Control*, vol. 4, pp. 567–576; 1958. (Dec 543)
Woodward, P. M., "Probability and Information Theory, with Applications to Radar," McGraw-Hill Book Co., Inc., New York, N. Y., and Pergamon Press, London, Eng.; 1953. (Jun 329, 330, 348, 361, 385, 397, 400, 412) (Dec 519)
Woodward, P. M., and Davies, I. L., "A theory of radar information," *Phil. Mag.*, ser. 7, vol. 41, pp. 1001–1017; October, 1950. (Sep 420)
Woodward, P. M., and Davies, I. L., "Information theory and inverse probability in telecommunication," *Proc. IEE*, vol. 99, pt. 3, pp. 37–44; March, 1952. (Jun 329, 367) (Dec 518)

Y

Youla, D. C., "The solution of a homogeneous Wiener-Hopf integral equation occurring in the expansion of second-order stationary random functions," IRE Trans. on Information Theory, vol. IT-3, pp. 187–193; September, 1957. (Sep 484)
Youla, D. C., "The use of maximum likelihood in estimating continuously modulated intelligence which has been corrupted by noise," IRE Trans. on Information Theory, vol. IT-3, pp. 90–105; March, 1954. (Mar 7) (Sep 420)

Z

Zadeh, L. A., "Optimum nonlinear filters," *J. Appl. Phys.*, vol. 24, pp. 396–404; April, 1953. (Mar 54) (Sep 420, 500)
Zadeh, L. A., and Ragazzini, J. R., "An extension of Wiener's theory of prediction," *J. Appl. Phys.*, vol. 21, pp. 645–655; July, 1950. (Sep 423, 426, 484)
Zadeh, L. A., and Ragazzini, J. R., "Optimum filters for the detection of signals in noise," Proc. IRE, vol. 40, pp. 1223–1231; October, 1952. (Jun 329, 349)
Zierler, N., "Legendre Sequences," Lincoln Lab., Mass. Inst. Tech., Lexington, Group Rept. No. 34–71; May 2, 1958. (Jun 329)
Zierler, N., "Linear recurring sequences," *J. Soc. Indus. Appl. Math.*, vol. 7, pp. 31–48; March, 1959. (Mar 18) (Jun 329) (Sep 463) (Dec 523)
Zierler, N., "On a Variation of the First-Order Reed-Muller Codes," Lincoln Lab., Mass. Inst. Tech., Lexington, Group Rept. No. 34–80; October, 1958. (Sep 465)
Zierler, N., "Several binary-sequence generators," *Proc. Am. Math. Soc.*, vol. 7, pp. 675–681; August, 1956. (Jun 329)
Zierler, N., "Several Binary-Sequence Generators," Lincoln Lab. Mass. Inst. Tech., Lexington, Tech. Rept. No. 95.

A STATEMENT OF EDITORIAL POLICY

The IRE TRANSACTIONS ON INFORMATION THEORY is a quarterly journal devoted to the publication of papers on the transmission, processing, and utilization of information. The exact subject matter of acceptable papers is intentionally, by editorial policy, not sharply delimited. Rather, it is hoped that as the focus of research activity changes, a flexible policy will permit the TRANSACTIONS to follow suit and that it will continue to serve its readers with timely articles on the fundamental nature of the communication process. Topics of current appropriateness include the coding and decoding of digital and analog communication transmissions, studies of random interferences and of information bearing signals, analyses and design of communication and detection systems, pattern recognition, learning, automata, and other forms of information processing systems.

Papers can be of two kinds, tutorial or research, and should be so indicated. The former must be well-written expositions summarizing the state of a field in which research is still in progress, or else unify results scattered in the literature. Research papers must be original contributions not published elsewhere. They must present new methods, concepts, or ideas, or extend old ones to new areas of applicability; or, they must present new data, findings or inventions, or solve new problems of more than casual interest. They will not be accepted if, in the view of the reviewers and editors, they constitute a straightforward and easy application of existing theory to a special case of limited interest. It is not necessary that the length of each research paper be great; on the contrary, the submission of short but formal research notes is to be encouraged.

In addition to papers, readers are invited to submit notes to the Correspondence section. These may include such things as early summaries of important work to be published later at greater length, or remarks on material that has already appeared. Contributions in the form of "problem statements" are also sought for the Correspondence section. This category includes problems to which the author knows no solution but suspects that another reader might, conjectures for which a proof or disproof is desired, and so forth.

INFORMATION FOR AUTHORS

Authors are requested to submit editorial correspondence or technical manuscripts to the Editor for possible publication in the PGIT TRANSACTIONS. Papers submitted should include a statement as to whether the material has been copyrighted, previously published, or submitted for publication elsewhere.

To expedite reviewing procedures, it is requested that authors submit the original and two legible copies of all written and illustrative material. The manuscript should be double-spaced, and the illustrations drawn in India ink on drawing paper or drafting cloth. Each paper should include a carefully written abstract of not more than 200 words. Papers should be prepared for publication in a manner similar to those intended for the PROCEEDINGS OF THE IRE. Further instructions may be obtained from the Editor. The original copy and drawings of material not accepted for publication will be returned.

All technical manuscripts and editorial correspondence should be addressed to Arthur Kohlenberg, Melpar, Inc., 11 Galen Street, Watertown 72, Mass.

Local Chapter activities and announcements, as well as other nontechnical news items, should be addressed to the PGIT Newsletter, c/o Prof. N. M. Abramson, Electrical Engineering Department, Stanford University, Stanford, Calif.

INSTITUTIONAL LISTINGS

The IRE Professional Group on Information Theory is grateful for the assistance given by the firms listed below and invites application for Institutional Listing from other firms interested in the field of Information Theory.

IBM RESEARCH, INTERNATIONAL BUSINESS MACHINES CORP., Yorktown Heights, N. Y.
Error Correcting & Detecting Codes, Theory of Assemblies & Automata, Information Networks, Reliability

REPUBLIC AVIATION CORP., Farmingdale, N. Y.
Aircraft, Missiles, Drones, Electronic Analyzers; U. S. Distr. of Alouette Turbine-Powered Helicopter

The charge for an Institutional Listing is $50 per issue or $150 for four consecutive issues. Applications for Institutional Listings and checks (made payable to the Institute of Radio Engineers) should be sent to L. G. Cumming, Institute of Radio Engineers, 1 East 79 St., New York 21, N. Y.